Inflationary stages in cosmological models with a scalar field

V. A. Belinskiĭ, L. P. Grishchuk, Ya. B. Zel'dovich, and I. M. Khalatnikov

L. D. Landau Institute of Theoretical Physics, USSR Academy of Sciences

(Submitted 14 March 1985)
Zh. Eksp. Teor. Fiz. **89**, 346–360 (August 1985)

The solutions of the gravitational equations for a homogeneous isotropic universe filled with a massive scalar field are investigated. Particular attention is paid to the question of the generality and conditions of realization of inflationary stages of expansion. It is shown that inflationary stages are an unavoidable property of a large class of solutions. The results of the paper indicate that the concept of spontaneous creation of the universe with a subsequent inflationary stage probably does not require the fulfillment of any too special requirements.

§1. INTRODUCTION

This paper analyzes all possible solutions of the gravitational equations for a homogeneous isotropic universe with scalar field having a definite rest mass. This investigation, which is of interest in its own right, is motivated by the ideas of so-called inflationary cosmological models. Although the possibility of an inflationary (de Sitter) stage in the evolution of the early universe was noted long ago (see, for example, the bibliography in the review Ref. 1), the possible part it could play in the solution of cosmological problems was most clearly pointed out in Ref. 2. It must be said that it is however now clear that by no means all variants of inflationary models can stand up to comparison with observational data. The necessary amplitude of perturbations cannot be readily obtained in models with specific Higgs dependence of the potential on the field. In addition, the very concept of a de Sitter stage, arising after the singularity and a hot stage of expansion, has lost its attraction compared with variants which include spontaneous quantum creation of the universe. It seems that the viable variant of the theory is currently the one in which it is assumed that in the early universe there existed a scalar field φ with values exceeding m_p, where $m_p = 1.22\times 10^{19}$ GeV is the Planck mass.[1)] It has been noted[1] that if in addition the initial value of $\dot{\varphi}$ is sufficiently small then the scale factor $a(t)$ of the homogeneous isotropic universe increases in accordance with a nearly exponential law, i.e., there is an inflationary (quasi-de Sitter) stage of expansion. However, the degree of generality of solutions possessing an inflationary stage; the possible occurrence of these stages for large initial values of $\dot{\varphi}$; the quantitative characterization of "advantageous" and "disadvantageous" cases; and the modifications in the set of possible solutions introduced by a nonvanishing spatial curvature—all these remain unresolved questions. To solve them, we use the methods of the qualitative theory of dynamical systems.

We consider the simplest case of massive (with mass m) minimally coupled scalar field described by the Lagrangian

$$L_\varphi = -{}^1/_2\varphi_{;k}\varphi^{;k} - {}^1/_2 m^2\varphi^2 \tag{1.1}$$

in a homogeneous isotropic cosmological model with metric

$$-ds^2 = -dt^2 + a^2(t)\,(dx_1^2+dx_2^2+dx_3^2) \times [1+{}^1/_4 k(x_1^2+x_2^2+x_3^2)]^{-1}, \tag{1.2}$$

where $k=1$, $k=-1$, and $k=0$ correspond, respectively, to closed, open, and flat models.

In this case, only the diagonal components of the energy-momentum tensor of the scalar field are nonzero, having the same form as for an ideal isotropic fluid with certain effective energy density ε and effective pressure p. Concretely,

$$T_0{}^0 = -\varepsilon, \quad T_\alpha{}^\beta = p\delta_\alpha{}^\beta,$$

where

$$\varepsilon = {}^1/_2\dot{\varphi}^2 + {}^1/_2 m^2\varphi^2, \quad p = {}^1/_2\dot{\varphi}^2 - {}^1/_2 m^2\varphi^2. \tag{1.3}$$

From these expressions it is clear that when $\dot{\varphi}^2 \ll m^2\varphi^2$ the effective equation of state is $p = -\varepsilon$, this indicating the possible occurrence of a quasi-de Sitter stage. In the opposite limiting case $\dot{\varphi}^2 \gg m^2\varphi^2$ and, in particular when $m=0$, we have the maximally hard equation of state $p=\varepsilon$. Cosmological solutions near the singularity for such an equation of state were considered in Ref. 3. The equation of state $p=\varepsilon$ itself was proposed earlier in Ref. 4 in a different physical realization. Finally, in the regime of an oscillating φ the averaged p vanishes, mimicking a dust medium. Thus, in the different regimes the homogeneous field $\varphi(t)$ possesses different effective equations of state. It is instructive to examine this paradoxical situation in the example of an idealized model in which it is assumed that the Hubble parameter, $H = \dot{a}/a$, is constant, the relation $|H| \gg m$ holding, and the "reaction" of the field φ on the scale factor $a(t)$ can be ignored.[5] For $H=$ const, the equation for φ [see Eq. (2.1) below] can be solved exactly. It can be shown[5] that in the case of expansion, i.e., for $H>0$, all solutions for the field φ (except one) behave in such a way that with the passage of time the equation of state tends to $p=-\varepsilon$. Conversely, in the case of contraction all solutions (except one) behave in such a way that the equation of state tends to $p=\varepsilon$. The qualitative features of this behavior are still present in the general case too.

We note that in the late stages of expansion there develop oscillations of the field φ describing "dust" matter consisting of scalar particles at rest. Their decay and transformation into hot plasma are not considered here. Also left on one side are important questions such as the possible role played by spatial inhomogeneity of the field φ and the metric.

We shall consider the solutions of the classical equa-

 0038-5646/85/080195-09$04.00

tions for $\varphi(t)$ and $a(t)$ right down to the singularities, which are a typical property of cosmological models (although some closed-model solutions do not have singularities at all). However, from the physical point of view the solutions of the classical equations are invalid at densities and curvatures which reach or exceed the Planck values. Nominally, the limit of applicability of the classical solutions can be taken to be the field values for which

$$\varepsilon = {}^1/_2\dot{\varphi}^2 + {}^1/_2 m^2\varphi^2 \approx m_p^4$$

(no further restrictions on $a(t)$ are required). The initial data for the classical stages of the evolution are determined at this limit. A particular set of initial values can be regarded as the consequences of the solution of the corresponding quantum-gravity problem in the region $\varepsilon \gtrsim m_p^4$. Here, we adhere to the idea advanced in Ref. 6,[2)] according to which the cosmological singularity must ultimately be replaced by "an act of spontaneous creation of the universe." It was shown in Ref. 6 that an inflationary stage of expansion is necessary if a universe created with characteristic Planck dimensions is to be able to grow in a definite time to macroscopic sizes and, ultimately, after transition to the hot plasma stage, to scales equal to or exceeding those of the currently observed universe. Putting aside for the moment the quantum-mechanical analysis and the discussion of the probability of occurrence of particular initial data, we elucidate the fate in the classical regime of solutions characterized by arbitrary initial values of φ and $\dot{\varphi}$. Of particular interest are closed models ($k = 1$), since it is precisely for them that the concept of "creation" is more meaningful.[3)]

The conclusions obtained in the present paper indicate that an inflationary stage is a very general property of the solutions considered and, thus, the concept of quantum "creation" of the universe with a subsequent inflationary stage probably does not require fulfillment of any too special conditions.

§2. BASIC EQUATIONS

The Lagrangian (1.1) and the metric (1.2) lead to the following simultaneous system of gravitational equations and equations for the field φ:

$$\ddot{\varphi} + 3H\dot{\varphi} + m^2\varphi = 0, \tag{2.1}$$

$$\dot{H} + H^2 = {}^1/_6 \varkappa m^2\varphi^2 - {}^1/_3 \varkappa \dot{\varphi}^2, \tag{2.2}$$

$$H^2 + ka^{-2} = {}^1/_6 \varkappa (m^2\varphi^2 + \dot{\varphi}^2), \tag{2.3}$$

where

$$\varkappa = 8\pi G = 8\pi m_p^{-2}, \quad H = \dot{a}/a. \tag{2.4}$$

Equations (2.1) and (2.2) form a three-dimensional dynamical system in the phase space φ, $\dot{\varphi}$, H. However, it is more convenient to use the variables φ, $\dot{\varphi}$, H, t together with the dimensionless quantities x, y, z, and η:

$$\varphi = 3sm^{-1}x, \quad \dot{\varphi} = 3sy, \quad H = mz, \quad t = m^{-1}\eta, \quad s = (12\pi)^{-1/2} m m_p. \tag{2.5}$$

Then Eqs. (2.1) and (2.2) become

$$x_\eta = y, \quad y_\eta = -x - 3zy, \quad z_\eta = x^2 - 2y^2 - z^2, \tag{2.6}$$

and the relation (2.3) can be written in the form

$$x^2 + y^2 - z^2 = km^{-2}a^{-2}. \tag{2.7}$$

Here and in what follows, the subscript η denotes the derivative with respect to this variable. The relation (2.4) for the Hubble parameter now takes the form

$$z = a_\eta / a. \tag{2.8}$$

In what follows, it will be convenient to present the results both in terms of x, y, z, η as well as in terms of φ, $\dot{\varphi}$, H, t. The transition from the one set of variables to the other is one to one and should not present difficulties.

Equations (2.6) do not contain the parameter k and thus describe all three models simultaneously in the phase space xyz. Equation (2.7) shows the regions of phase space in which the trajectories of the various models lie. First of all, it is clear that the surface of the cone $x^2 + y^2 - z^2 = 0$, which corresponds to the flat case $k = 0$, separates the regions containing the trajectories of the open model $k = -1$ (interior of the cone, which contains the z axis) from the region containing the trajectories of the closed model $k = 1$ (the part of the phase space exterior to the cone and containing the plane xy). The trajectories of the flat model lie on the cone itself and, therefore, form a two-dimensional invariant phase space.

We note that the sections of the trajectories lying in the upper half of the phase space ($z > 0$) correspond to expansion of the model (i.e., $H > 0$) and those in the lower ($z < 0$) to contraction ($H < 0$). At the same time, the trajectories of the open and flat models cannot be continued from the upper half to the lower. They are separated by a singular point, the coordinate origin $(x, y, z) = (0, 0, 0)$, the apex of the upper and lower sheets of the cone. Physically, this point (if approached from the region $H > 0$) corresponds to the final stages of unlimited expansion ($a \to \infty$), which terminate the evolution of the models with $k = -1$ and $k = 0$. For $k = 1$, the trajectories can intersect the plane $z = 0$ ($H = 0$), this corresponding to the times of the extrema of the scale function ($\dot{a} = 0$).

Thus, the expansion and contraction for the models with $k = -1$ and $k = 0$ are to be considered separately. However, all the trajectories describing the contraction in these models can be obtained from the trajectories corresponding to expansion by two symmetry transformations that the system of equations (2.6)–(2.8) possesses:

$$\begin{aligned} &\text{a)}\ \eta \to -\eta, && \text{b)}\ \eta \to -\eta, \\ &\quad z \to -z, && \quad z \to -z, \\ &\quad x \to -x, && \quad y \to -y. \end{aligned} \tag{2.9}$$

We now turn to the investigation of the simplest case: $k = 0$.

§3. FLAT MODEL

For $k = 0$, the system (2.6)–(2.7) simplifies. We shall consider only an expanding model, for which $H > 0$. Substituting $z = +(x^2 + y^2)^{1/2}$ in the second equation in (2.6), we obtain a two-dimensional dynamical system in the variables x, y:

$$x_\eta = y, \quad y_\eta = -x - 3y(x^2 + y^2)^{1/2}. \tag{3.1}$$

As phase space of this system, we shall use the xy plane, though it must be remembered that the true trajectories of the system (3.1) lie on the cone $z = +(x^2+y^2)^{1/2}$ and the phase diagram in the xy plane is the conceptual orthogonal projection of the true picture onto the horizontal plane. It is readily seen that in a finite region of variation of x and y the system (3.1) has only one singular point—the origin $(x, y) = (0, 0)$. A simple analysis shows that this point is a stable focus, and the asymptotic behavior of a solution near it has the form

$$x=\frac{2}{3\eta}\sin(\eta-\eta_0),\quad y=\frac{2}{3\eta}\cos(\eta-\eta_0),$$

$$\eta_0=\text{const},\quad \eta\to+\infty. \qquad (3.2)$$

For the variable z we have at the same time

$$z=2/3\eta. \qquad (3.3)$$

Thus, as we have already said, the point $(x, y) = (0, 0)$ corresponds to the final stages of unlimited expansion, in which the field φ oscillates, being damped, and the scale factor tends to infinity in accordance with the law

$$a\propto\eta^{2/3}. \qquad (3.4)$$

All the remaining singular points of the system (3.1) lie at infinity, and to study them one can use a standard device such as mapping of the xy plane (taking it to be in the horizontal position) onto the lower half of a sphere of unit radius lying on this plane and touching it at the origin (so-called Poincaré sphere). The points of the plane are mapped onto the sphere by means of central projection (from the center of the sphere). The infinitely distant points of the plane are then projected onto the equator of the sphere. If we now project the entire lower half of the sphere back onto the horizontal plane, but this time by means of vertical projection, we obtain as a result a continuous and one-to-one mapping of the xy plane onto the interior of the disk of unit radius. Investigation of the nature of the singular points on the boundary of this disk completely describes the infinity of the phase space.

If we make all the analysis associated with this procedure, we arrive at the results illustrated in Fig. 1. Figure 1a shows the behavior of the trajectories in the xy plane, Fig. 1b the mapping of this phase diagram onto the disk of unit radius as described above. On the boundary of the disk ($x^2+y^2=\infty$) the system has four singular points: two repelling nodes, K_1 and K_2, and two saddles, S_1 and S_2. At the center of the disk (i.e., at the coordinate origin $x = 0, y = 0$) there is the attracting focus F with asymptotic behavior (3.2) with which we are already acquainted. All the trajectories emanate from the four infinitely distant singular points and then wind around the central focus. We write down the asymptotic behavior of the solutions to the system (2.1)–(2.3) near the singular points at infinity. To the point K_1 there corresponds an initial cosmological singularity at a certain finite time, which can always be chosen at $t = 0$. Near K_1, we have

$$\varphi=sm^{-1}\ln(t/t_0),\qquad H=1/3t, \qquad (3.5)$$

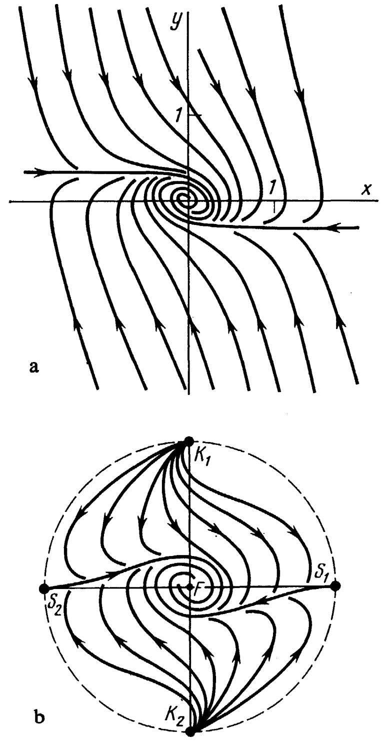

FIG. 1.

where $t_0 > 0$ is an arbitrary constant. Emergence from K_1 corresponds to increasing t, beginning from $t = 0$. In the neighborhood of this node, the effective equation of state is $p = \varepsilon$.

Near the saddle S_1 the asymptotic behavior of the singular solution corresponding to the emerging separatrix S_1F describes cosmological evolution of a different type, which begins in the infinitely distant past:

$$\varphi=-st,\qquad H=-{}^1/_3m^2t. \qquad (3.6)$$

Here, the time increases from the value $t = -\infty$, corresponding to the initial singularity. Near this separatrix, in the region of sufficiently large φ, the effective equation of state is $p = -\varepsilon$.

We note that the asymptotic value of $\dot\varphi$ in (3.6) is

$$\dot\varphi=-s=-(12\pi)^{-1/2}mm_p. \qquad (3.7)$$

This means that in the φ, $\dot\varphi$ plane the separatrix S_1F has a horizontal asymptote (3.7), to which S_1F tends as $\varphi\to\infty$. The asymptotic behaviors near the two other singular points K_2 and S_2 and the properties of the separatrix S_2F are analogous and follow from the obvious symmetry properties.

Since the effective equation of state $p = -\varepsilon$ is realized near the separatrix S_1F for sufficiently large values of φ (determined more precisely below), it is in this region that we must expect the occurrence of inflationary stages.

We now turn to analytic construction of solutions near this separatrix. As can be seen from (1.3), the equation of state $p = -\varepsilon$ is realized (for positive φ) in the region

$$|\dot\varphi|\ll m\varphi. \qquad (3.8)$$

This is the region in which the separatrix S_1F passes. As follows from Eqs. (2.2) and (2.3), the condition (3.8) in the case $k=0$ is equivalent to the requirement

$$|\dot{H}|\ll H^2. \qquad (3.9)$$

Taking into account (3.8), we obtain from (2.3) for $k=0$

$$H\approx(m^2/3s)\varphi. \qquad (3.10)$$

Substituting this expression for H in Eq. (2.1), we reduce it to the form

$$\frac{d\dot{\varphi}}{d\varphi}=-\frac{m^2}{s}\frac{\varphi(\dot{\varphi}+s)}{\dot{\varphi}},$$

whence we obtain the following approximate equation for the phase trajectories:

$$(\dot{\varphi}+s)\exp(-\dot{\varphi}/s)=C\exp(m^2\varphi^2/2s^2), \qquad (3.11)$$

where C is an arbitrary constant. From the same equations (2.1) and (3.10) we can obtain the ratio of the scale factor at some initial time t_i to the scale factor at some final time t_f. It follows from these equations that

$$H=(d\dot{\varphi}/dt)/3(\dot{\varphi}+s),$$

and since

$$\frac{a(t_f)}{a(t_i)}=\exp\int_{t_i}^{t_f}H(t)\,dt,$$

we obtain

$$a(t_f)/a(t_i)=|[\dot{\varphi}(t_i)+s]/[\dot{\varphi}(t_f)+s]|^{1/3}. \qquad (3.12)$$

It can be seen from this that appreciable growth of the scale factor is possible only for the trajectories that at the time t_f approach sufficiently close to the separatrix S_1F [see (3.6), (3.7)].

It is well known that for cosmological applications the ratio $a(t_f)/a(t_i)$ for the inflationary stage must be of order 10^{30} (or greater). It is readily seen that a ratio of this order can be obtained in the sections of the phase trajectories that begin with $|\dot{\varphi}+s|\sim mm_p$ and end with $|\dot{\varphi}+s|\sim10^{-90}mm_p$ (or even smaller values of $|\dot{\varphi}+s|$). With regard to the sections of the trajectories in which $|\dot{\varphi}|$ varies from the largest possible values $\sim m\varphi$, situated at the limit of the region of applicability of these approximate solutions, to values corresponding to $|\dot{\varphi}+s|\sim mm_p$, we find that $a(t_f)/a(t_i)$ on these sections does not exceed $(m_p/m)^{1/3}$. Indeed, this follows from Eq. (3.12) if for $|\dot{\varphi}(t_i)|$ we take the maximal possible value $\sim m\varphi(t_i)\sim m_p^2$. Since additional cosmological arguments associated with the growth of small perturbations and the bound on the anisotropy of the background radiation[4] require $m/m_p\sim10^{-5}$–10^{-6} (see the discussion below), growth of the scale factor by $(m_p/m)^{1/3}$ times is indeed insignificant compared with what is needed. One can also show that outside the region of applicability of the approximate solution (3.11)–(3.12), i.e., in the sections of the trajectories where $\dot{\varphi}^2\gg m^2\varphi^2$, the growth of the scale factor is also slight, and the ratio $a(t_f)/a(t_i)$ is determined by the same quantity $(m_p/m)^{1/3}$.

The ratio $a(t_f)/a(t_i)$ can be expressed in terms of the initial and final values of the field φ itself. The expression (3.12) with allowance for (3.11) and the approximate equality

$$\exp(-\dot{\varphi}/s)\approx1,$$

which is clearly valid in the range of variation of $|\dot{\varphi}|$ in which we are interested (i.e., for $|\dot{\varphi}+s|\lesssim mm_p$), can be rewritten in the form

$$a(t_f)/a(t_i)=\exp[2\pi m_p^{-2}(\varphi_i^2-\varphi_f^2)], \qquad (3.13)$$

where φ_i and φ_f are, respectively, the initial and final values of the field φ. If as φ_f we take the value $\varphi_f\sim m_p$, at which the final stages of the expansion with damped oscillations of the field φ begin, to ensure realization of the necessary inflationary stage (i.e., one in which $a(t_f)/a(t_i)\sim10^{30}$) it is necessary to take $\varphi_i\sim(3$–$4)m_p$ (see also Ref. 1). But if the initial $\varphi_i\gg m_p$, then the duration of the necessary inflationary stage in terms of $\Delta\varphi=\varphi_i-\varphi_f$ is determined by the condition

$$\Delta\varphi\approx6m_p^2/\varphi_i.$$

In particular, for φ_i taken on the boundary of the quantum domain we have $\Delta\varphi\sim m\ll m_p$.

We emphasize that we were above discussing the duration (with respect to the variable φ) of the minimal necessary inflationary stage, characterized by inflation by 10^{30} times. In fact, the actually realized duration of the inflationary stage in the solutions of the considered model is almost always greater and is determined by an initial value of the field φ_i for which the trajectory has approached sufficiently close to the separatrix S_1F. In the sequel, such a trajectory can only continue to approach the separatrix, not move away from it. For this reason, the end of the inflationary stage for any trajectory will always correspond to the value $\varphi_f\sim m_p$, i.e., to the value after which the separatrix passes into the concluding noninflationary regime of expansion with oscillations. Thus, for $\varphi_i\sim(3$–$4)m_p$ the duration of the inflationary stage that arises corresponds to the minimal necessary. But if $\varphi_i\gg m_p$, the degree of inflation in such a solution can actually greatly exceed 10^{30}. Thus, inflation from $\varphi_i\sim m_p^2/m$ (near the quantum limit) to $\varphi_f\sim m_p$ leads to

$$a(t_f)/a(t_i)\sim\exp(m_p^2/m^2).$$

In this case, only a minute portion of the entire volume of the universe is accessible to contemporary observations, that part in fact which passed through the inflationary stage at the very end.

As already noted, bounds on the growth of small perturbations and the anisotropy of the background radiation lead to additional conditions on the parameters of inflationary models. In particular, $H(t)$ at the end of the inflationary stage must not be too large, namely, of order $(10^{-5}$–$10^{-6})m_p$. Since the end of such a stage always corresponds in the model considered to the value $\varphi_f\sim m_p$, it follows from (3.10) that $H(t_f)\sim m$, and it is this that leads to the estimate mentioned above for the mass of the scalar field: $m\sim(10^{-5}$–$10^{-6})m_p$.

All the analysis is of course also valid (after the substitution $\varphi\to-\varphi$, $\dot{\varphi}\to-\dot{\varphi}$) for the region near the separatrix S_2F.

We now turn to the question of the degree of generality of solutions possessing an inflationary stage. This problem is most readily solved by means of the exactly constructed

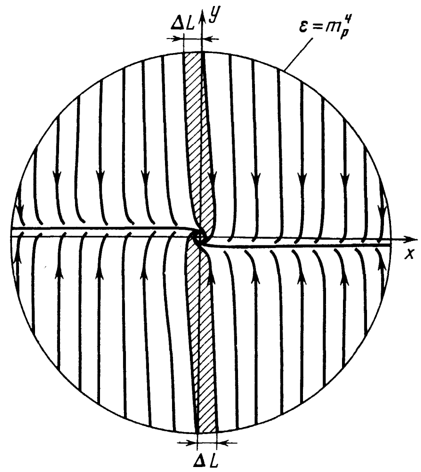

FIG. 2.

phase diagram of the system (3.1) in the region within the quantum limit on which we specify the initial data. If this construction is carried out, then the phase diagram on large scales (out to the quantum boundary) can be represented qualitatively as shown in Fig. 2. It can be seen from this figure that the phase trajectories are practically vertical outside the central region with radius $\sim m_p$. In regions with $|\varphi| > m_p$, the trajectories turn sharply near the separatrices and pass along them right to the central region, where they then wind round the focus $x = 0, y = 0$.

As already noted, the advantageous trajectories (i.e., those possessing the required inflationary stage) are all the ones for which $\varphi_i > (3\text{–}4)m_p$. These trajectories begin at the points of the quantum boundary situated everywhere on the circle except its two segments near the y axis (see Fig. 2). The set of disadvantageous trajectories, i.e., the ones that begin at the points of these segments, is small, since the length of each such segment is $\Delta L \sim (6\text{–}8)m_p$, whereas the length of the complete quantum circle is $L \sim 2\pi m_p^2/m$. Thus, the ratio $\Delta L/L$, which determines the measure of the disadvantageous trajectories, is

$$\Delta L/L \approx m/m_p \ll 1.$$

Thus, an inflationary stage, which begins sooner or later, is inherent in almost all trajectories that begin on the quantum limit. This conclusion is very important, since it indicates a great generality of inflationary regimes in models with $k = 0$.

An analogous conclusion also holds for the models we investigated of a scalar field with self-interaction proportional to $\lambda\varphi^4$, and also with Higgs potential. We shall not dwell on the details but merely give the phase diagrams corresponding to these cases.

Figure 3a shows the phase diagram for the flat model for which the scalar field is described by the Lagrangian

$$L_\varphi = -{}^1/_2\varphi_{;m}\varphi^{;m} - {}^1/_2 m^2\varphi^2 - {}^1/_4\lambda\varphi^4.$$

We have used the same dimensionless variables (2.5) as for $\lambda = 0$. To be specific, we assume $\lambda \sim 10^{-12}$, $m \sim 10^{-6} m_p$. The asymptotic equation of the inflationary separatrices at large $|x|$ is now

$$y = -\gamma x, \quad \gamma = (\lambda/6\pi)^{1/2} m_p m^{-1}.$$

Figure 3b shows the phase diagram for the flat model with Higgs scalar field. Its Lagrangian is

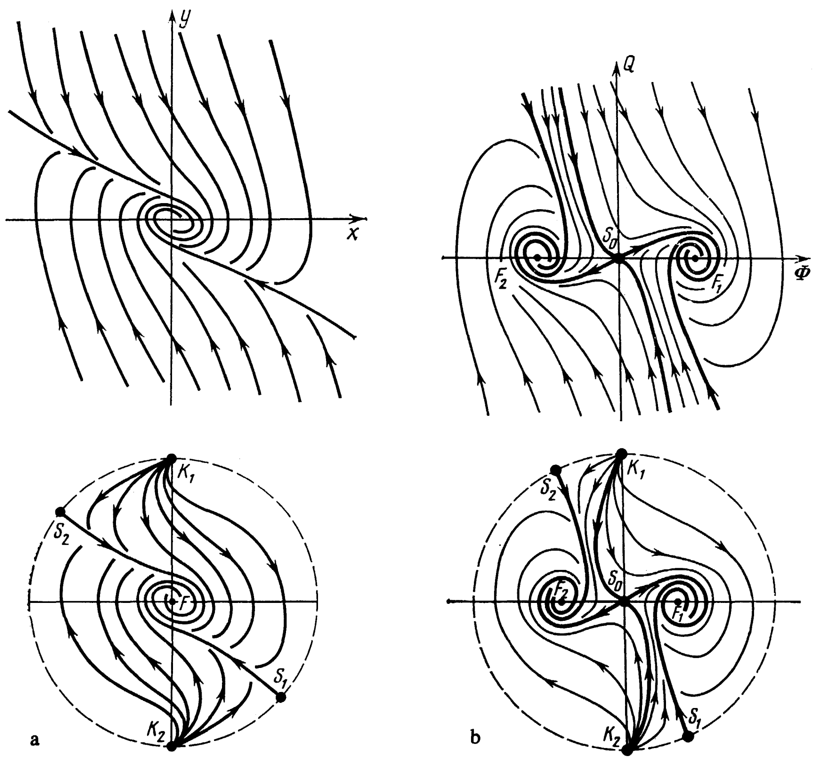

FIG. 3.

$$L_\varphi = -{}^1/_2\varphi_{;m}\varphi^{;m} - {}^1/_4\lambda(\varphi^2-\varphi_0^2)^2,$$

where to be specific we assume $\lambda \sim 10^{-12}$ and $\varphi_0 \sim 3\times 10^{-2} m_p$. As phase variables we have chosen the dimensionless quantities

$$\Phi = \varphi_0^{-1}\varphi, \quad Q = (9\lambda/2)^{-1/2}\varphi_0^{-2}\dot{\varphi}.$$

In this case there are two focuses F_1 and F_2 (for $\Phi = \pm 1$) and a saddle S_0 in the center between them. All the separatrices are shown by the heavy lines. The asymptotic equation of the inflationary separatrices S_1F_1 and S_2F_2 at large $|\Phi|$ is

$$Q = -\beta\Phi, \quad \beta = (27\pi)^{-1/2}\varphi_0^{-1}m_p.$$

§4. BEHAVIOR OF THE TRAJECTORIES AT THE PHASE SPACE INFINITY

First of all, we complete the description begun in Sec. 2 of the three-dimensional phase space of the system (2.6). It can be seen from (2.6) that in a finite region of variation of x, y, z this system has just one equilibrium state—the origin $(x, y, z) = (0, 0, 0)$. Other singular points can lie only at infinity. To investigate their number and nature, it is convenient to compactify the phase space and complete it by the infinitely distant boundary $x^2+y^2+z^2 = \infty$, going over from the Cartesian coordinates x, y, z to the spherical

$$x = r\sin\theta\cos\psi, \quad y = r\sin\theta\sin\psi, \quad z = r\cos\theta$$

with subsequent transformation of the radius in accordance with the law $r = \rho(1-\rho)^{-1}$, where $0 \leqslant \rho \leqslant 1$. If we now introduce a new time τ from the condition $d\tau/d\eta = (1-\rho)^{-1}$, then in the variables ρ, θ, ψ, τ the system (2.6) finally takes the form

$$\rho_\tau = -\rho^2(1-\rho)\cos\theta(6\sin^2\theta\sin^2\psi + \cos 2\theta),$$
$$\theta_\tau = \rho\sin\theta\cos 2\theta(1-3\sin^2\psi), \qquad (4.1)$$
$$\psi_\tau = -(1-\rho) - 3\rho\cos\theta\sin\psi\cos\psi.$$

For $0 \leqslant \rho < 1$, Eqs. (2.6) are equivalent to the system (4.1), though the latter admits smooth continuation to the boundary $\rho = 1$. Thus, in the variables $0 \leqslant \rho \leqslant 1$, $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \psi \leqslant 2\pi$ the phase space of the system (2.6), augmented by the infinitely distant boundary, is compact and can be imagined as a ball of unit radius placed in the same phase space xyz (or rather, on its second copy) with center at the origin. Each point of the original phase space is mapped along its radius vector to a certain point in the interior of the ball, while the points $x^2+y^2+z^2 = \infty$ at infinity are mapped to the surface $\rho = 1$ of the ball. It is readily seen from (4.1) that on this surface the system has 14 singular points. They are all shown together with the interior of the ball in Fig. 4. Essentially different are only four: one point from each group P, K, S, and C. The properties of all the others are then obtained from the chosen four by some combination of the symmetry transformations (2.9).

At the north pole of the ball there is a saddle point $P(\theta = 0)$, from which there emanates into the phase space of the open model a two-dimensional pencil of trajectories (two-dimensional separatrix). Near P, these solutions have the asymptotic behavior

$$\varphi = \varphi_0 - {}^1/_8 m^2\varphi_0 t^2, \quad H = 1/t, \qquad (4.2)$$

where φ_0 is an arbitrary constant. The initial singularity $t = 0$ corresponds to the point P itself and emergence from it corresponds to increase of the time from this value. Near the P, the scale factor follows the law

$$a \propto t. \qquad (4.3)$$

Among these solutions there is a pure vacuum solution: $\varphi \equiv 0$ (for $\varphi_0 = 0$), whose trajectory is represented by the polar axis PF. It corresponds to the metric of a flat (four dimensionally) world with hyperbolic spatial sections (so-called Milne model).

The parallel $\theta = \pi/4$ is the intersection of the surface of the ball with the upper half of the cone $x^2+y^2-z^2 = 0$, i.e., maps the infinitely distant boundary of the two-dimensional phase space of the flat model about which we have already learnt (Sec. 3). On this parallel lie the saddles $S_1(\psi = 0)$ and $S_2(\psi = \pi)$, from which just one trajectory from each enters the ball. These two trajectories pass along the surface of the cone and are the flat-model separatrices S_1F and S_2F considered in Sec. 3. On the same parallel are the points $K_1(\psi = \pi/2)$ and $K_2(\psi = 3\pi/2)$, which are repelling nodes and send trajectories both into the space of the flat model (along the cone) as well as into the phase space of the open and closed models. These nodes correspond to initial cosmological singularities, and the asymptotic behavior near them does not depend on the model type. For K_1 we have the same asymptotic behavior (3.5), and for K_2 the same expressions with the substitution $\varphi \to -\varphi$.

On the equator $\theta = \pi/2$ of the surface of the ball there are four more singular points C, which are saddles. The points $C_1(\sin\psi = -1/\sqrt{3})$ and $C_2(\sin\psi = 1/\sqrt{3})$ are the infinitely distant ends of the straight line lying in the intersection of the planes $z = 1/\sqrt{2}$ and $x = -\sqrt{2}y$. This line is an asymptote for the two-dimensional pencils of trajectories that emanate from C_1 and C_2 into the phase space of the closed model. These trajectories (like the separatrices S_1F

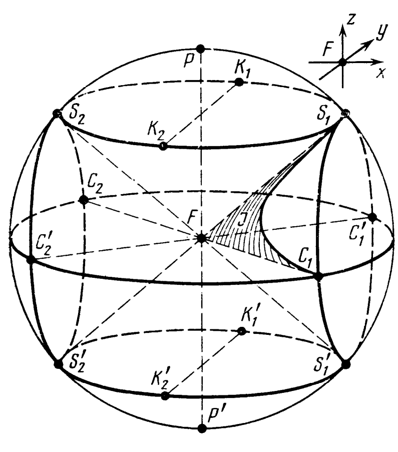

FIG. 4.

and S_2F) describe cosmological evolution that began in the infinitely distant past. The points C_1 and C_2 themselves correspond to the time $t = -\infty$ (initial singularity), and emergence from them corresponds to increase in the time from this value. The leading terms in the asymptotic behavior of these solutions near C_1 and C_2 are

$$\varphi = \text{const}\cdot\exp(-mt/\sqrt{2}), \quad H = m/\sqrt{2}, \tag{4.4}$$

where the constant in front of the exponential is positive for C_1 and negative for C_2. The scale factor here increases from zero exponentially rapidly:

$$a \propto \exp(mt/\sqrt{2}), \tag{4.5}$$

but this does not however mean that in this region an inflationary stage of interest to us arises; for it follows from (1.3) that near C_1 and C_2 the effective equation of state is $\varepsilon + 3p = 0$ and

$$\varepsilon \approx {}^3/_4 m^2\varphi^2 \propto \exp(-\sqrt{2}mt) \propto a^{-2}.$$

Thus, the energy density decreases with increasing t too rapidly for the requirements that we usually impose on the concept of inflationary stages to be satisfied.

The saddle points C_1' ($\sin\psi = 1/\sqrt{3}$) and C_2' ($\sin\psi = -1/\sqrt{3}$) (the ends of the line that is the intersection of the planes $z = -1/\sqrt{2}$ and $x = \sqrt{2}y$) relate to the lower half of the phase space (contraction) and attract corresponding two-dimensional pencils.

On the lower half of the surface of the ball (see Fig. 4) there are also the singular points P', K', and S', symmetric to the points P, K, S. The south pole $P'(\theta = \pi)$ attracts a two-dimensional pencil of trajectories, including the purely vacuum FP'. The nodes K_1' ($\theta = 3\pi/4$, $\psi = \pi/2$) and K_2' ($\theta = 3\pi/4$, $\psi = 3\pi/2$) also attract trajectories and correspond to final cosmological singularities. The asymptotic behavior near these collapse points again does not depend on the type of the model. The saddles S_1' ($\theta = 3\pi/4$, $\psi = 0$) and S_2' ($\theta = 3\pi/4$, $\psi = \pi$) are each approached by one trajectory, the separatrices FS_1' and FS_2' of the contracting flat model.

We have described above the nature of the singular points with respect to the three-dimensional physical phase space of the system (2.6), the interior of the ball $0 \leqslant \rho < 1$. For complete description of the phase infinity it is necessary to add to these results the diagram of all the unphysical trajectories that lie entirely on the surface $\rho = 1$ of the ball. By continuity arguments, such trajectories determine the behavior of the physical integral curves of the system that lie within the ball but near its surface. Setting $\rho \equiv 1$ in (4.1), we obtain

$$\theta_\tau = \sin\theta\cos 2\theta(1 - 3\sin^2\psi), \qquad \psi_\tau = -3\cos\theta\sin\psi\cos\psi. \tag{4.6}$$

The phase diagram for this two-dimensional dynamical system is represented in Fig. 5 as if the angles θ and ψ were Cartesian coordinates. To cover with them the surface of the ball shown in Fig. 4, it is necessary to identify the left-hand edge $\psi = 0$ with the right-hand $\psi = 2\pi$, identify with one another all points of the upper edge $\theta = 0$ and also identify with one another all the points of the lower edge $\theta = \pi$. With respect to the two-dimensional space $\theta\psi$, all the points K and S are simple nodes, while the remainder are simple saddles. All the saddle separatrices are shown in Fig. 5 by heavy lines. For covenience of matching Figs. 5 and 4, the polar angle θ in Fig. 4 is plotted downward.

We note that the system (4.6) can also be integrated explicitly. It phase trajectories, shown in Fig. 5, are described by the equation

$$\sin^{2/3}\psi\cos^{4/3}\psi\sin^2\theta\cos^{-1}2\theta = \text{const}. \tag{4.7}$$

§5. OPEN MODEL

We discuss briefly the solutions for an expanding open model. Their compact phase diagram is bounded by the surface of the cone $z = +(x^2+y^2)^{1/2}$ and the section of the spherical surface including the north pole P (see Fig. 4).

The quantum boundary is now shown by the surface of the cylinder $x^2 + y^2 = 8\pi m_p^2/3m^2$ (i.e., $\varepsilon = m_p^4$), which intersects the surface of the cone along the circle that is the quantum boundary of the flat models and which we discussed earlier (see Fig. 2). For the nonflat models, the boundary surface is two dimensional, since, compared with the flat case, it is necessary to specify one further independent parameter—the initial value of H (or a).

It is obvious that the trajectories which begin on the surface of the cylinder but near its intersection with the cone

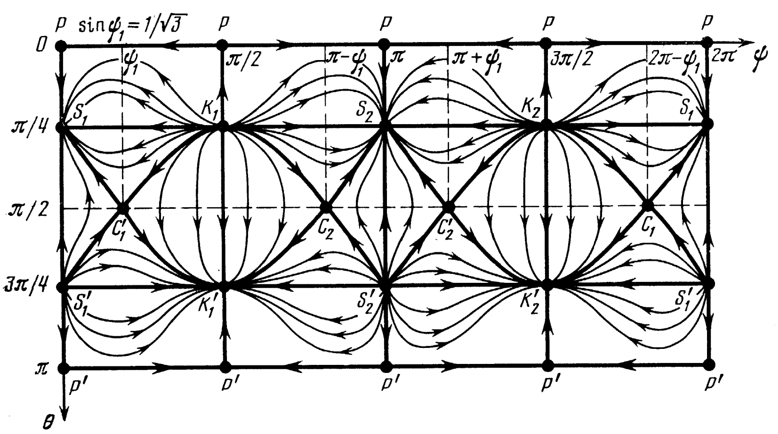

FIG. 5.

will have qualitatively the same properties as the trajectories of the flat model, i.e., the majority of them approach and then pass along the separatrices S_1F and S_2F, which lie on the surface of the cone, and are subject to a long inflationary stage. These trajectories are characterized by the fact that the spatial curvature, already small in the region in which the initial data are specified, decreases even further with the passage of time in the inflationary stage. With regard to the remaining trajectories which intersect the quantum boundary far from this circle, their fate is less certain. It can however be seen that some of them approach the cone in the regions of the separatrices S_1F and S_2F within the classical region and undergo an inflationary stage. But others reach the region of oscillations near the focus F without having been inflated.

It is important to note that in all solutions the spatial curvature, even in the ones in which it has decreased strongly in the inflationary stage, ultimately becomes dynamically important in the final stages of unlimited expansion in the region of oscillations near the focus F (effective equation of state $p = 0$). This is reflected in the fact that all the trajectories approaching the focus approach ever closer to the vertical vacuum line PF.

Thus, an inflationary stage is an unavoidable intermediate stage for a set of solutions in the open model, although it is now not so easy to give a numerical estimate of the corresponding measure as it was in the case of the flat model.

§6. CLOSED MODEL

As already noted, the complete phase space for the case $k = 1$ is the interior of the ball apart from the interior regions of the cone $x^2 + y^2 - z^2 = 0$, in which the trajectories of the open model lie. Compared with the cases $k = 0$ and $k = -1$, there is a new possibility—the trajectories can intersect the plane $H = 0$ $(z = 0)$, i.e., there are points of regular maxima or minima of the scale factor $a(t)$. These points are strictly separated. For their description, we find first the surface in the phase space on which $\dot{H} = 0$ $(z_\eta = 0)$. As can be seen from the last relation of (2.6), this surface is the cone with equation $x^2 - 2y^2 - z^2 = 0$. As shown in Fig. 4, the sheets of this cone are arranged horizontally (containing within them the x axis) and intersect the infinitely distant boundary of the phase space along the curves $S_1C_1S'_1C'_1$ and $S_2C_2S'_2C'_2$. These curves intersect the equator of the surface of the ball at the singular points C, and the complete cone intersects the horizontal plane $z = 0$ along the two straight lines $C'_1FC'_2$ and C_1FC_2, which have on it the equations $x = \sqrt{2}y$ and $x = -\sqrt{2}y$, respectively. The cone lies entirely in the region of space of the closed model and touches the phase surface of the flat model (i.e., the cone $x^2 + y^2 - z^2 = 0$) along the two generators $z = \pm x$. The lines $x = \pm\sqrt{2}y$ on the plane $z = 0$ are the ones that separate the points at which a maximum or a minimum of the scale factor is attained. Every trajectory which intersects the plane $z = 0$ in the region $x^2 > 2y^2$ is a solution possessing a regular minimum of $a(t)$. The trajectories that intersect the plane $z = 0$ at points of the region $x^2 < 2y^2$ have at these points a maximum of $a(t)$. On every trajectory lying within the cone $\dot{H} = 0$ the Hubble parameter H increases with the time t, i.e., motion in this direction is directed only upward. Outside this cone, H decreases, and the motion along the trajectories is directed downward.

To establish the degree of generality of the inflationary stages in the closed model, it is important to establish the points of the cone $\dot{H} = 0$ at which the trajectories can leave the cone. Finding on the surface of this cone the curve at whose points $\ddot{H} = 0$, one can show that in the expansion phase $(H > 0)$ and for positive values of φ emergence from the cone is possible only through a narrow region on its surface, denoted by the letter J and hatched in Fig. 4. For negative φ, an analogous region exists on the left-hand sheet of the cone in the triangle S_2C_2F. At large values of $|\varphi|$, the ones in which we are mainly interested, the trajectories leaving the cone immediately enter the neighborhood of the separatrices S_1F and S_2F, and the corresponding solutions undergo a prolonged inflationary stage. Since the region of emergence from the cone adjoins the separatrices S_1F and S_2F, this means that almost all trajectories that leave the cone enter into an inflationary regime. This applies to all trajectories that appear in the expansion phase already within the cone $\dot{H} = 0$, i.e., from points of a regular minimum of the scale factor (we take no account of the previous history of such solutions) and from the singular points C_1 and C_2. Some of the trajectories that begin at the singularities K enter an inflationary regime in the same way, i.e., by passing first through the interior region of the cone $\dot{H} = 0$. There are trajectories that make the transitions $K_1 \to S_1F$ and $K_2 \to S_2F$. The other trajectories which begin at K approach the separatrices S_1F and S_2F without passing into the cone $\dot{H} = 0$. These are the trajectories, for example, that pass near the trajectories of the flat model along the paths $K_1 \to S_2F$ or $K_2 \to S_1F$. It is difficult to estimate quantitatively the relative number of solutions possessing the required inflationary stage, but it can be seen from our discussion that his number is fairly large.

Besides the solutions considered, there is a set of trajectories along which it is possible to pass directly (avoiding the region near the separatrices S_1F and S_2F) from the initial singularity K to the collapse K', this being done by following a path either through one expansion maximum or through a finite number of oscillations of the scale factor between regular minima and maxima.

The presence of a closed chain of trajectories lying on the boundary of the phase space of the closed model (for example, $S_1FS'_2S_2FS'_1S_1$) indicates the possible existence of periodic solutions, which are represented by closed trajectories near this chain. The realization of such solutions requires, of course, a very special choice of the initial data, since on the section FS'_2, moving, for example, in the direction of the saddle point S'_2, the trajectory must pass for a long time along the unstable solution FS'_2. Solutions of such kind were discussed earlier in Ref. 12.

We must point out the existence of solutions that are t-symmetric with respect to the moments of the regular maxima or minima of $a(t)$. One can show that these trajectories intersect the plane $z = 0$ at points of the y or the x axis,

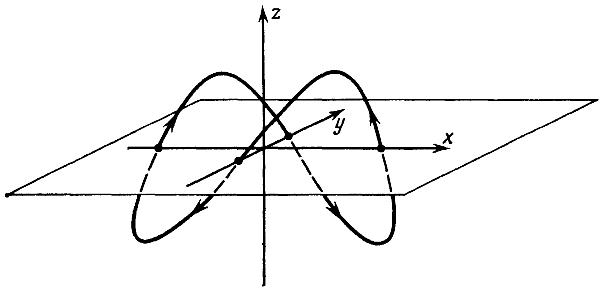

FIG. 6.

respectively. Also possible is the existence of periodic solutions t-symmetric both with respect to the moments at which a_{max} is attained as well as the moments corresponding to a_{min} (Fig. 6). The possible existence of such solutions was noted in Ref. 13, and the question of the existence of infinitely oscillating nonperiodic solutions is discussed in Ref. 14. As noted in Ref. 13, the periodic solutions form an infinite discrete set. Under such conditions, it is possible to have trajectories one of whose ends begins or terminates at the singularities K and K' while the other winds around and closer and closer to one of the periodic solutions. Also not impossible is the existence of trajectories containing a finite number of revolutions around the periodic solutions, passing successively from one of them to another. The existence in a dynamical system of classes of trajectories with the properties described above usually leads to stochastization of the regime of its behavior.

We note that although the considered system of equations admits an infinite number of solutions corresponding to an "eternally oscillating universe" these solutions cannot serve as the basis for any realistic model of the universe without singularities. First, even in the framework of the idealized system we have considered these solutions require special specification of the initial data and in this sense are degenerate. Second, allowance for physical mechanisms such as decay of the oscillations of the field φ, growth of inhomogeneities in the contraction stage, quantum production of particles, growth of entropy, etc., will undoubtedly reduce (most probably to zero) the already small probability of realization of such solutions.

[1)] In the paper, we use a system of units in which the velocity of light, Planck's constant, and Boltzmann's constant are equal to unity. Latin indices take the values 0, 1, 2, 3, Greek the values 1, 2, 3. The interval is written in the form $-ds^2 = g_{ik} dx^i dx^k$, where g_{ik} has the signature $(-+++)$. For time, we use the notation $x^0 = t$. Differentiation with respect to t is denoted by a dot.

[2)] See also the pioneering studies of Refs. 7 and 8.

[3)] There is already a fairly extensive literature devoted to questions of the quantum creation of the universe (see Ref. 9 and the references given there).

[4)] For the restrictions on the parameters of inflationary models, see, for example, Ref. 10 and the recent Ref. 11. In the case of quasi-de Sitter models, which are ones we consider, these restrictions apply to the final stages of the inflation.

[1] A. Linde, Rep. Prog. Phys. **47**, 925 (1984).
[2] A. Guth, Phys. Rev. D **23**, 347 (1981).
[3] V. A. Belinskiĭ and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **63**, 1121 (1972) [Sov. Phys. JETP **36**, 591 (1973)].
[4] Ya. B. Zel'dovich, Zh. Eksp. Teor. Fiz. **41**, 1609 (1961) [Sov. Phys. JETP **14**, 1143 (1962)].
[5] Ya. B. Zeldovich, in: *Soviet Scientific Reviews, Section E: Astrophysics and Space Physics Reviews* (ed. R. A. Syunyaev), Harwood Academic Publishers (1985).
[6] L. P. Grishchuk and Ya. B. Zeldovich, in: *Quantum Structure of Space and Time* (eds. M. Duff and C. Isham), Cambridge University Press (1982), p. 409.
[7] E. Tryon, Nature **246**, 396 (1973).
[8] P. I. Fomin, Dokl. Akad. Nauk SSSR **9A**, 831 (1975).
[9] A. D. Linde, Zh. Eksp. Teor. Fiz. **87**, 369 (1984) [Sov. Phys. JETP **60**, 211 (1984)].
[10] V. A. Rubakov, M. V. Sazhin, and A. V. Verjaskin, Phys. Lett. **B115**, 189 (1982).
[11] S. W. Hawking, Phys. Lett. **B150**, 339 (1985).
[12] A. A. Starobinskii, Pis'ma Astron. Zh. **4**, 155 (1978) [Sov. Astron. Lett. **4**, 82 (1978)].
[13] S. W. Hawking, Lectures Given at the Les Houches Summer School, July (1983).
[14] D. Page, Classical and Quantum Gravity **1**, 417 (1984).

Translated by Julian B. Barbour