ON THE 90TH ANNIVERSARY OF L D LANDAU'S BIRTH

Singularity, initial conditions and quantum tunneling in modern cosmology

I M Khalatnikov, A Yu Kamenshchik

Contents

1. Introduction	525
2. Oscillatory approach to singularity	526
3. Quantum tunneling and wave function of the Universe, model with complex scalar field	528
4. Energy scale of inflation and loop effects	532
5. Hamiltonian quantization of gravitation and quantum self-consistency	534
6. Dynamics of cosmological models with scalar field: chaos and fractality	536
7. Perturbation methods in cosmology	539
8. Conclusions	540
References	540

<u>Abstract.</u> The key problems of modern cosmology, such as the cosmological singularity, initial conditions, and the quantum tunneling hypothesis, are discussed. The relationship between the latest cosmological trends and L D Landau's old ideas is analyzed. Particular attention is given to the oscillatory approach to singularity; quantum tunneling processes determining wave function of the Universe in the presence of a complex scalar field; and the role of quantum corrections in these processes. The classical dynamics of closed models with a real scalar field is investigated from the standpoint of chaotic, fractal, and singularity-avoiding properties.

1. Introduction

Talking with his students in the 1950s, Lev Davidovich Landau listed three problems as the most important for theoretical physics still waiting to be solved: the problem of the cosmological singularity, the problem of phase transitions, and the problem of superconductivity [1]. Today one may say that there has been a major breakthrough in the studies of superconductivity [2] and phase transitions [3]. At the same time, the problem of the cosmological singularity still remains intriguing, in spite of the substantial advances in studies of different aspects of it.

For example, R Penrose and S Hawking [4] proved the impossibility of indefinite continuation of geodesics under certain conditions. This was interpreted as pointing to the

I M Khalatnikov, A Yu Kamenshchik L D Landau Institute of Theoretical Physics, Russian Academy of Sciences, ul. Kosygina 2, 117334 Moscow, Russia Tel. (7-095) 137 32 44 Fax (7-095) 938 20 77 E-mail: khalat@itp.ac.ru, kamen@landau.ac.ru

Received 5 March 1998 Uspekhi Fizicheskikh Nauk **168** (6) 593–611 (1998) Translated by A S Dobroslavskiĭ; edited by L V Semenova existence of a singularity in the general solution of the Einstein equations. However, these theorems were based on topological methods, and did not allow for finding the particular analytical nature of the singularity. It was only in the papers by E M Lifshitz and I M Khalatnikov [5, 6], and V A Belinskii, E M Lifshitz and I M Khalatnikov [7, 8, 12] that the analytical behavior of the general solutions of the Einstein equations was investigated in the neighborhood of the singularity. It is important that the results exhibited a dependence on the required number of arbitrary functions, which was an indication of their universality [9, 10]. These papers for the first time, revealed the intriguing phenomenon of an oscillatory approach of the Universe to the singularity [5-13], which has become known as the *Mixmaster Universe* [14]. A simple model with three degrees of freedom (a Bianchi IX model of the Universe) was used to demonstrate that the Universe approaches the singularity in such a way that its contraction along two axes is accompanied by an expansion with respect to the third axis, and the axes change their roles according to a rather complicated law. The dynamics of the oscillatory approach to the singularity turned out to be extremely rich, and attracted the attention of researchers using a broad variety of mathematical methods, from number theory [15] to the theory of catastrophes [16].

From a physical standpoint, we are especially interested in the analysis of the probability distribution of the initial data for the expanding oscillatory Universe, started in Ref. [11] and further developed in Ref. [13]. This analysis anticipated the quantum-cosmological attempts at establishing the baseline for cosmological evolution through constructing the wave function of the Universe [17, 18]. It would be appropriate here to recall Landau's words [19] that a consistent physical theory of the future should include not only the correct equations of motion, but also the initial conditions. It is this problem that is one of the main tasks of quantum cosmology today.

It is interesting that the concept of quantum tunneling has recently been widely used for constructing the wave function 23], which is discussed in the textbook on non-relativistic quantum mechanics by L D Landau and E M Lifshitz [24] with references to Landau's paper from as early as 1932! Another feature of the quantum birth of the Universe owing to quantum tunneling and the change of sign of space-time is that this allows the avoidance of the 'fall into singularity'. In other words, the models of quantum birth of the Universe assume that, considering the inverse time evolution of the Universe, there is a recoil rather than a singularity — the collapse of the Universe stops short and expansion begins. The evolution towards a smaller Universe is only possible through an analytical extension of time into the complex plane — in other words, through tunneling. In this way, the wave function of the Universe based on the concept of quantum tunneling allows avoidance of the singularity, which is intuitively attractive. It is interesting that the study of classical cosmological scenarios avoiding the singularity goes well back in time [25-29] and is still being pursued [30-29]33].

As regards the idea of Landau [19] that a consistent theory should define not only the correct equations of motion but also the initial conditions, it is naturally related to another idea which has recently become very popular. This is what one might call the concept of quantum self-consistency. It appears that the quantum theory of elementary particles and fundamental forces should define the content of the matter fields and hence the spectrum of particles, while the relevant restrictions can be derived from the requirement of intrinsic self-consistency in the theory. A classic example is the socalled mechanism of Glashow-Iliopoulos-Maiani [34], who introduced a fourth quark into the theory of electroweak interactions in order to suppress the neutral currents affecting the strangeness and to eliminate the chiral anomaly. Later this quark was observed experimentally and became known as 'charmed'. Even more spectacular was the appearance of the critical dimensions in the theories of strings and superstrings [35], which pose a claim to becoming fundamental theories of physics. As is known, the theory of boson strings can be consistently formulated in a space of dimensionality d = 26[36], whereas a superstring exists in a space of dimensionality d = 10.

It is well known that cosmological observations may serve as a test ground for various models of the physics of elementary particles (see, for example, Ref. [37]). A less known fact is that even the requirement of self-consistency of quantum cosmology imposes various restrictions on the spectrum of the theory [38-44, 22, 23].

In this review we are going to consider some directions in contemporary quantum and classical cosmology which relate to such ideas of Landau as the importance of the problem of the singularity, the complexification of time associated with tunneling, the relevance of the initial conditions in the framework of the fundamental theory, and the related concept of quantum self-consistency. Section 2 is devoted to the oscillatory approach to singularity; in Section 3 we consider the theory of cosmological tunneling transitions using the model with a complex scalar field [45–49, 32, 33]. Section 4 deals with the calculation of loop corrections in quantum cosmology, and with such aspects of the wave function of the Universe as its normalizability and ability to

predict the initial conditions of cosmological evolution. In Section 5 we consider the scheme of the Hamiltonian BRST (Becchie–Rue–Stora–Tyutin) quantization of closed cosmological models and its implications. Sections 6 and 7 deal with the feasibility of avoiding the singularity even within the framework of classical cosmology with a scalar field. The main results of these studies are summarized in the conclusion.

2. Oscillatory approach to the singularity

One of the first exact solutions obtained within the framework of the general theory of relativity was the Kasner solution [50] for a cosmological model like Bianchi I, which considers a gravitation field in empty space with Euclidean metrics whose time dependence is described by

$$ds^{2} = dt^{2} - t^{2p_{1}} dx^{2} - t^{2p_{2}} dy^{2} - t^{2p_{3}} dz^{2}, \qquad (2.1)$$

where the exponents p_1, p_2, p_3 satisfy the condition

$$p_1 + p_2 + p_3 = p_1^2 + p_2^2 + p_3^2 = 1.$$
 (2.2)

Arranging the exponents in the order

$$p_1 < p_2 < p_3$$
, (2.3)

one can parametrize them:

$$p_{1} = \frac{-u}{1+u+u^{2}}, \qquad p_{2} = \frac{1+u}{1+u+u^{2}},$$

$$p_{3} = \frac{u(1+u)}{1+u+u^{2}}.$$
(2.4)

If the parameter *u* varies in the range $u \ge 1$, the exponents p_1 , p_2 , p_3 all assume allowable values:

$$-\frac{1}{3} \leq p_1 \leq 0, \quad 0 \leq p_2 \leq \frac{2}{3}, \quad \frac{2}{3} \leq p_3 \leq 1.$$
 (2.5)

With u < 1, we have the same ranges for p_1, p_2, p_3 , since

$$p_1\left(\frac{1}{u}\right) = p_1(u), \quad p_2\left(\frac{1}{u}\right) = p_3(u), \quad p_3\left(\frac{1}{u}\right) = p_2(u).$$
(2.6)

For Bianchi VIII and Bianchi IX models, the Kasner regime (2.1), (2.2) is no longer an exact solution of the Einstein equations; it is possible, however, to construct the generalized Kasner solutions [5-8]. It is also possible to construct a perturbation theory for which the exact Kasner solution (2.1), (2.2) serves as the zero approximation, while the perturbations are represented by those terms in the Einstein equations which depend on the curvature of space (such terms are obviously not present in the Bianchi I model). This perturbation theory holds in the neighborhood of the singularity — that is, at $t \rightarrow 0$. The remarkable fact is that these perturbations lead to the transition from one Kasner regime to another with different parameters. The metrics of the generalized Kasner solution in the synchronous frame can be written as

$$ds^{2} = dt^{2} - (a^{2}l_{\alpha}l_{\beta} + b^{2}m_{\alpha}m_{\beta} + c^{2}n_{\alpha}n_{\beta}) dx^{\alpha} dx^{\beta}, \quad (2.7)$$

where

$$a = t^{p_l}, \quad b = t^{p_m}, \quad c = t^{p_n}.$$
 (2.8)

The three-dimensional vectors **l**, **m**, **n** define the directions in which the spatial dimensions exhibit a power-law time dependence (2.8). Let $p_l = p_1$, $p_m = p_2$, $p_n = p_3$, so that

$$a \sim t^{p_1}, \quad b \sim t^{p_2}, \quad c \sim t^{p_3},$$
 (2.9)

that is, the Universe contracts in the directions of **m** and **n**, and expands along **l**.

It was demonstrated [7] that the perturbations caused by the terms with spatial curvature lead to a transition to a different Kasner regime characterized by the following formulas:

$$a \sim t^{p'_l}, \quad b \sim t^{p'_2}, \quad c \sim t^{p'_3},$$
 (2.10)

where

$$p_{1}' = \frac{|p_{1}|}{1 - 2|p_{1}|}, \quad p_{m}' = -\frac{2|p_{1}| - p_{2}}{1 - 2|p_{1}|},$$
$$p_{n}' = -\frac{p_{3} - 2|p_{1}|}{1 - 2|p_{1}|}.$$
(2.11)

We see that the perturbations give rise to a transition from one 'Kasner epoch' to another, in which the negative power of t shifts from direction **l** to direction **m**. In the course of transition, the function a(t) achieves its maximum, and the function b(t) goes through a minimum. Accordingly, the previously decreasing b starts to increase, and a starts to decrease, whereas c(t) remains a decreasing function. The perturbation that initiated the transition from regime (2.9) to regime (2.10) decreases and almost disappears. Then another perturbation starts to grow and gives rise to a transition from this Kasner epoch to yet another one, and so on.

We would like to emphasize that the successful application of perturbation theory relies on the fact that the perturbation affects the dynamics of the system in such a way that the initial perturbation is eliminated. A rather remote but interesting analogy can be traced with the phenomenon of asymptotic freedom in quantum field theory [51], where the perturbative calculation of quantum corrections in non-Abelian Yang–Mills model and some other models reduces the magnitude of the relevant coupling constants, thus justifying the method itself.

Going back to the rules that govern the shift of the negative exponent from one direction to another, one can show that they can be expressed with the aid of a parametrization of Eqn (2.4):

$$p_l = p_1(u), \quad p_m = p_2(u), \quad p_n = p_3(u)$$
 (2.12)

and then

$$p'_{l} = p_{2}(u-1), \quad p'_{m} = p_{1}(u-1), \quad p'_{n} = p_{3}(u-1).$$

(2.13)

The larger of the two positive exponents remains positive.

Consecutive changes of Eqn (2.13), accompanied by the toggle of the negative exponent between directions \mathbf{l} and \mathbf{m} , continue till the integer part of u is exhausted — that is, until u

becomes less than one. Then, in accordance with Eqn (2.6), the value of u < 1 can be converted into u > 1. At this time one of the exponents p_l or p_m is negative, and p_n becomes the least of the two positive numbers ($p_n = p_2$). The next series of changes will toggle the negative exponent between directions **n** and **l**, or **n** and **m**.

Evolution of the model towards the singularity consists of consecutive periods (called eras), within which the dimensions along two axes oscillate, whereas the scale along the third axis decreases monotonically, and the volume decreases with time according to a near-linear law. In the course of the transition from one era to the next, the axes along which the scale decreases monotonically exchange their roles. The order of change of the eras of varying duration assume a random nature.

Each sth era corresponds to a decreasing sequence of values of the parameter u. This sequence has the form $u_{\max}^{(s)}, u_{\max}^{(s)} - 1, \dots, u_{\min}^{(s)}$, where $u_{\min}^{(s)} < 1$. Let us introduce the following notation:

$$u_{\min}^{(s)} = x^{(s)}, \qquad u_{\max}^{(s)} = k^{(s)} + x^{(s)},$$
 (2.14)

that is, $k^{(s)} = [u_{\max}^{(s)}]$ (the brackets indicate the largest integer less or equal to $u_{\max}^{(s)}$). The number $k^{(s)}$ determines the length of the era. For the next era we get

$$u_{\max}^{(s+1)} = \frac{1}{x^{(s)}}, \quad k^{(s+1)} = \left[\frac{1}{x^{(s)}}\right].$$
 (2.15)

The ordering of durations $k^{(s)}$ of consecutive eras (determined by the number of Kasner epochs in each) asymptotically tends to assume a random nature. The stochastic nature of the process is explained by rules (2.14), (2.15) which govern the transitions from one era to the next in the infinite series of values of u. If this sequence starts with a certain initial value $u_{\text{max}}^{(0)} = k^{(0)} + x^{(0)}$, then the lengths of sequences $k^{(0)}, k^{(1)}, \ldots$ can be represented as the numbers involved in the expansion of the continued fraction

$$k^{(0)} + x^{(0)} = k^{(0)} + \frac{1}{k^{(1)} + 1/(k^{(2)} + ...)}$$
 (2.16)

This succession of eras can be described statistically if in place of the given initial value $u_{max}^{(0)} = k^{(0)} + x^{(0)}$ we consider the distribution $x^{(0)}$ over the interval (0, 1), governed by a certain probability law. Then it is possible to obtain certain distributions of the values $x^{(s)}$, truncating each *s*th set of numbers. It can be demonstrated that, as *s* increases, these distributions tend to a stationary (independent of *s*) probability distribution w(x), for which $x^{(s)}$ is completely 'forgotten':

$$w(x) = \frac{1}{(1+x)\ln 2} \,. \tag{2.17}$$

From Eqn (2.17) it follows that the probability distribution of the lengths of series of k is given by

$$W(k) = \frac{1}{\ln 2} \ln \frac{(k+1)^2}{k(k+2)}.$$
 (2.18)

Moreover, it is also possible to calculate exactly the probability distributions for other parameters, such as the parameter δ which defines the relationship between the

amplitudes of logarithms of functions a, b, c and the logarithmic time [13].

In this way, a statistical analysis of the evolution in the neighborhood of the singularity [11] indicates that the stochasticity and probabilistic distributions of parameters already arise in classical general relativity. One may say that this probabilistic description of the classical Universe at the beginning of its existence may serve as prototype of the probability distribution for the initial state of the Universe which arises in quantum cosmology as a result of construction of the wave function of the Universe.

The dynamics of an oscillatory approach to singularity are still being actively discussed in the literature. For example, the fact that the dynamics are actually chaotic was challenged on the grounds that the corresponding Lyapunov exponents vanish [52]. Later, however, it was found that different parametrizations of the time variable give rise to different values of the Lyapunov exponents. This called for the development of invariant methods for studying chaos in cosmology. The equations which describe the approach to singularity using the Painlevé test were analyzed — that is, the search for singular points whose location depends not only on the equations themselves, but also on the initial conditions. The type of singularity at these points allows the possible nonintegrability of the dynamic equations to be detected [53]. The results of these studies point to the chaotic nature of the approach to the singularity [54], although the authors point out that these statements do not yet hold the status of theorems.

In Ref. [15], the study of chaos was based on observerindependent fractal techniques. A large number of quasiperiodical paths were constructed — that is, paths for which the durations of the Kasner eras obey a periodical law, where the total volume of the Universe monotonically decreases. It is interesting that these quasi-periodical paths correspond to the so-called periodical irrational values of the parameter *u*, represented by periodical continued fractions. It was known even to Lagrange that such numbers are the roots of quadratic equations with integer coefficients[†]. The set of these numbers is countable and has measure zero over the set of all values of the parameter u. These paths form a strange repeller of non-trivial fractal dimension. The existence of such a repeller and its parameters brought the authors of Ref. [15] to the definite conclusion that the oscillatory approach to singularity has a chaotic nature.

Another interesting aspect of the approach to singularity is the question concerning the composition of the matter that preserves the oscillations. It has long been known [10] that the assumption of the negligibly small role of matter near the singularity only holds for matter whose equation of state satisfies the condition $p \leq 2\epsilon/3$, where p is the pressure and ϵ is the density of energy. At the same time, it has been demonstrated [55] that the general solution of the gravitation equations for the cosmological model with a massless scalar field exhibits monotonic power-law asymptotic behavior instead of the oscillatory regime. This is quite natural, since this model is equivalent to the model with hydrodynamic matter whose equation of state is $p = \epsilon$. In the light of this fact it would be quite natural to conclude that the oscillatory

[†] Finite continued fractions correspond to rational numbers. After a number of oscillations started at a certain rational value of parameter u, the Universe arrives at the state with u = 1, whose metric is equivalent to the metric of the Minkowski space up to the Lorentz transform [9].

regime is not realized in string cosmology [56], since the massless dilaton field is an integral part of string cosmological models [57].

The nature of the dynamic evolution in the Bianchi IX model was also studied [58] using the very popular formalism of Ashtekar variables [59]. Using them in the context of the Hamiltonian formalism, it was demonstrated that the Bianchi II models may be treated as perturbations of the Bianchi I models, and are integrable. The Bianchi IX model is in turn a perturbation of the Bianchi II model; in this case, however, the algorithm of integrability does not work, which in the opinion of the authors explains the chaotic dynamics of the Bianchi IX model.

To conclude this section, we shall mention paper [16] where the dynamics of the Bianchi IX model were analyzed using the methods of the theory of catastrophes [60]; then the transition from one Kasner era to the next is equivalent to a fold-type catastrophe.

In this way, the study of the oscillatory approach to singularity attracts much interest, and is in a sense becoming a chapter of modern mathematical physics in its own right.

3. Quantum tunneling and wave function of the Universe, model with a complex scalar field

In recent years, the branch of theoretical physics known as quantum cosmology has been on the rise. It is important because it is linked with the development of a consistent scheme of quantization of gravitation, which is necessary for constructing a unified theory of fundamental interactions, as well as because it has the potential for predicting the initial conditions for the inflation model of the Universe [61, 37], which has won the status of a verifiable and even observable theory after the discovery of the anisotropy in the background radiation [62] predicted almost two decades ago [63].

Inflation cosmology is essentially based on the assumption that the early cosmological evolution passed through an inflational (or quasi-De-Sitter) stage of near-exponential expansion, which was then superseded by the conventional Friedmann power-law expansion. This assumption removes a number of old difficulties from the theory of a hot Universe and correctly accounts for the anisotropy of the background radiation.

The question concerning the origin of the inflationary Universe, however, remains open. To resolve it, one has to address the principles of quantum cosmology. The main principle of quantum cosmology is that the Universe is treated as a single quantum object, whereas the quantization of gravitation is performed in the language of dynamics of systems with constraints. Such an approach goes back to the early works of P Dirac, J A Wheeler, B S DeWitt, and others [64]. In the context of this approach, one may define the wave function of a closed Universe which obeys the Wheeler – DeWitt equation

$$\mathcal{H}\Psi(^{(3)}g,\Phi) = 0.$$
 (3.1)

Here $\Psi({}^{(3)}g, \Phi)$ is the wave function of the Universe, which depends on the 3-metric ${}^{(3)}g$ and the matter fields Φ , whereas \mathcal{H} is a constraint of the first kind called the super-Hamiltonian, whereby all fields and their canonically conjugate momenta are interpreted as operators satisfying the canonical commutative relations.

The most common approach to the Wheeler-DeWitt equation is based on considering the so-called mini-superspaces which only involve a finite number of cosmological degrees of freedom, the rest being 'frozen'. Another useful trick in studying Eqn (3.1) consists in the semiclassical expansion of the functional integral formally representing the solution of the Wheeler - DeWitt equation. This approach has been actively pursued since the early 1980s, when two ideologically close prescriptions for constructing the wave function of the Universe were proposed [17, 18]. Both approaches draw an analogy between the birth of the Universe 'from nothing' and the processes of tunneling in quantum mechanics and quantum field theory, treated in the semiclassical approximation. The probabilities of such processes can be evaluated by calculating the Euclidean action on the instantons — the solutions of the Euclidean equations of motion for the systems in question [65].

On the tree level — that is, in the lowest order of the perturbation theory — these functions can be represented in the following form:

(1) the so-called 'no-boundary' wave function [17]

$$\Psi_{\rm NB} \sim \exp(-I) \,, \tag{3.2}$$

(2) the tunneling wave function [18]

$$\Psi_{\rm T} \sim \exp(+I)$$
. (3.3)

Here *I* is the Euclidean action, which in the case of gravitation occurs with a minus sign in the equations of motion, as opposed to the situation with the conventional field theory. One can demonstrate that both wave functions (3.2) and (3.3) satisfy the Wheeler–DeWitt equation (3.1) [21, 39]. The proof of this fact relies heavily on the absence of a time dependence in Eqn (3.1). Apparently, the probability distributions obtained from the wave functions (3.2) and (3.3) are complementary to each other in the sense that the maximum of the no-boundary wave function corresponds to the minimum of the tunneling function, and vice versa.

One may say that the problem of reconciliation of quantum cosmology and inflation cosmology consists essentially in the search for a cosmological model whose wave function defines such initial conditions for the inflation stage as would correspond to today's large-scale structure. Most of the cosmological models of recent time involve the so-called inflaton scalar field with a non-zero classical mean, which ensures the existence of the effective cosmological constant at the early stage of cosmological evolution. Moreover, the simultaneous evolution of the space-time geometry and the inflaton scalar field ensures not only the inflation stage itself, but also a smooth transition to the Friedmann evolution.

Unfortunately, the simple cosmological models with a real scalar field, whose wave functions are calculated in the tree-level approximation in accordance with Eqn (3.2) or (3.3), are not capable of yielding a normalizable wave function of the Universe [66] and ensuring such a probability distribution for the initial value of the inflaton scalar field as would guarantee a sufficiently long inflation stage [67].

A way out of this situation may be sought in various directions. The one-loop treatment of the wave function of the Universe [38-43, 22, 23] gives a normalizable wave function for the Universe and an acceptable probability distribution for the initial value of the inflaton scalar field

provided that the spectrum of the theory is properly selected (see Section 4).

Another direction of development of quantum cosmology involves models more sophisticated than the model with a real scalar field. For example, in Refs [45-47] a model was considered with a complex scalar field. For one thing, complex scalar fields and non-Abelian multiplets of scalar fields arise naturally in contemporary theories of elementary particles. On the other hand, with a complex scalar field one has the opportunity to introduce an additional quasifundamental constant: the classical charge related to the global symmetry of the Lagrangian of the theory with respect to phase rotation of this field. The inclusion of this charge gives rise to a centrifugal term in the effective Hamiltonian, and the concept of a Euclidean region in the mini-superspace is changed drastically. Euclidean, or 'classically forbidden' regions become compact and host the instanton solutions with zero velocities on the boundary of the region [68]. Such instanton solutions correspond to an extremum of the Euclidean action and a peak in the probability distribution for the no-boundary wave function [47]. In this way, such instantons predict the initial conditions for the inflation stage, and the subsequent evolution of the model ensures a smooth transition to Friedmann's power-law expansion.

Recently, the models with non-minimal coupling between the inflaton scalar field and gravitation have become very popular [69]. On the one hand, such models offer many opportunities for comparing theoretical predictions with actual observations; on the other hand, they are more consistent from the standpoint of quantum gravitation [70]. The cosmological model with non-minimal coupling of the complex scalar field is quite rewarding [48, 49, 32, 33]. As indicated above, the model with minimal coupling of the scalar field admits only one instanton solution [45, 46]. This solution may be extended into the Lorentz region in accordance with the Lorentz equations of motion, thus providing for the origin of inflation. Such a scheme obviously favors a no-boundary wave function for the Universe [47]. At the same time, in the case of non-minimal coupling of the complex field we get a pair of instantons, of which one ensures a peak for the no-boundary wave function, whereas the other ensures a peak for the tunneling wave function.

Let us consider the model with non-minimal coupling of the complex scalar field in greater detail. Its action has the form

$$S = \int d^{4}x \sqrt{-g} \left[\frac{m_{\rm P}^{2}}{16\pi} (R - 2\Lambda) + \frac{1}{2} g^{\mu\nu} \phi_{\mu}^{*} \phi_{\nu} + \frac{1}{2} \xi R \phi \phi^{*} - \frac{1}{2} m^{2} \phi \phi^{*} - \frac{1}{4!} \lambda (\phi \phi^{*})^{2} \right], \qquad (3.4)$$

where ξ is the parameter of non-minimal coupling (for convenience, here we use a sign opposite to that commonly found in the literature), λ is the parameter of self-interaction of the scalar field, Λ is the cosmological term, *m* is the mass of the scalar field. The complex scalar field ϕ can be represented in the form

$$\phi = x \exp i\theta, \qquad (3.5)$$

where x is the absolute value of the complex scalar field, and θ is its phase. We consider the mini-superspace model with spatially homogeneous variables a (the cosmological radius in

the Friedmann–Robertson–Walker metric), x and θ . Now the action (3.4) becomes

$$S = 2\pi^{2} \int dt \, Na^{3} \left\{ \frac{m_{\rm P}^{2}}{16\pi} \left[6 \left(\frac{\dot{a}^{2}}{N^{2}a^{2}} + \frac{\ddot{a}}{N^{2}a} + \frac{1}{a^{2}} \right) - 2\Lambda \right] + \frac{1}{2N^{2}} \dot{x}^{2} + \frac{1}{2N^{2}} x^{2} \dot{\theta}^{2} + 3\xi \left(\frac{\dot{a}^{2}}{N^{2}a^{2}} + \frac{\ddot{a}}{N^{2}a} + \frac{1}{a^{2}} \right) x^{2} - \frac{1}{2} m^{2} x^{2} - \frac{1}{4!} \lambda x^{4} \right\},$$
(3.6)

where N is the lapse function. Integrating by parts, we eliminate the terms containing the second derivatives \ddot{a} , and rewrite Eqn (3.6) in a more convenient form:

$$S = 2\pi^{2} \int dt N \left\{ \frac{m_{\rm P}^{2}}{16\pi} \left[6 \left(-\frac{\dot{a}^{2}a}{N^{2}} + a \right) - 2\Lambda a^{3} \right] + \frac{1}{2N^{2}} \dot{x}^{2} \right. \\ \left. + \frac{1}{2N^{2}} x^{2} \dot{\theta}^{2} + 3\xi \left(\frac{-\dot{a}^{2}a}{N^{2}} + a \right) x^{2} - 6\xi \frac{\dot{a}\dot{x}a^{2}x}{N^{2}} \right. \\ \left. + \frac{1}{2N^{2}} \dot{x}^{2} a^{3} - \frac{1}{2} m^{2} x^{2} a^{3} - \frac{1}{4!} \lambda x^{4} a^{3} \right\}.$$
(3.7)

Observe that the phase variable θ is cyclical, and its canonically conjugate momentum p_{θ} must be conserved. We call its value the charge of the Universe and denote it by Q:

$$p_{\theta} = Q = a^3 x^2 \dot{\theta} \,. \tag{3.8}$$

Going over to the Hamiltonian formalism, we may rewrite the action (3.7) in the form

$$S = 2\pi^2 \int \mathrm{d}t \left(p_a \dot{a} + p_x \dot{x} - N\mathcal{H} \right), \qquad (3.9)$$

where the super-Hamiltonian \mathcal{H} is

$$\mathcal{H} = -\frac{p_a^2}{24a[m_P^2/(16\pi) + \xi x^2/2 + 3\xi^2 x^2]} - \frac{\xi p_x p_a x}{2a^2[m_P^2/(16\pi) + \xi x^2/2 + 3\xi^2 x^2]} + \frac{p_x^2}{2a^3} \frac{m_P^2/(16\pi) + \xi x^2/2}{m_P^2/(16\pi) + \xi x^2/2 + 3\xi^2 x^2} - U(a, x). \quad (3.10)$$

The function U(a, x), which will be referred to as the superpotential, is

$$U(a, x) = a \left[\frac{m_{\rm P}^2}{16\pi} (6 - 2\Lambda a^2) + 3\xi x^2 - \frac{Q^2}{a^4 x^2} - \frac{1}{2} m^2 x^2 a^2 - \frac{1}{24} \lambda x^4 a^2 \right].$$
 (3.11)

Variation of the action with respect to the lapse function N yields the constraint

$$\mathcal{H} = 0, \qquad (3.12)$$

whose quantum counterpart is the Wheeler - DeWitt equation (3.1).

Now it will be convenient to write down the effective Lagrangian which will only depend on a and x, and their derivatives (the lapse function N is selected to be unity):

$$L = \left\{ \frac{m_{\rm P}^2}{16\pi} \left[6(-\dot{a}^2 a + a) - 2\Lambda a^3 \right] + 3\xi(-\dot{a}^2 a + a)x^2 - 6\xi\dot{a}\dot{x}a^2x + \frac{1}{2}\dot{x}^2a^3 - \frac{Q^2}{2a^3x^2} - \frac{1}{2}m^2x^2a^3 - \frac{1}{4!}\lambda x^4a^3 \right\}.$$
(3.13)

This Lagrangian gives us the equations of motion

$$\frac{m_{\rm P}^2}{16\pi} \left(\ddot{a} + \frac{\dot{a}^2}{2a} + \frac{1}{2a} - \frac{\Lambda a}{2} \right) + \frac{\xi \dot{a}^2 x^2}{4a} + \frac{\xi \ddot{a} x^2}{2} + \xi x \dot{x} \dot{a} + \frac{\xi \dot{x}^2 a}{2} + \frac{\xi x^2 a}{2} + \frac{\xi x^2 a}{4a} + \frac{a \dot{x}^2}{8} - \frac{m^2 x^2 a}{8} + \frac{Q^2}{4a^5 x^2} - \frac{\lambda x^4 a}{96} = 0 \quad (3.14)$$

and

$$\ddot{x} + \frac{3\dot{x}\dot{a}}{a} - \frac{6\xi x\ddot{a}}{a} - \frac{6\xi\dot{a}^2 x}{a^2} - \frac{6\xi x}{a^2} + m^2 x - \frac{2Q^2}{a^6 x^3} + \frac{\lambda x^3}{6} = 0.$$
(3.15)

In addition, we can write the first integral of motion for our dynamic system:

$$-\frac{3}{8\pi}m_{\rm P}^2a\dot{a}^2 - 3\xi a\dot{a}^2x^2 - 6\xi x\dot{x}\dot{a}a^2 + \frac{a^3}{2}\dot{x}^2 - U(a,x) = 0.$$
(3.16)

Similar Euclidean equations of motion can be derived from Eqns (3.14)-(3.16) by changing the sign of the derivative-containing terms. Now we must define the concept of 'Euclidean regions' in our mini-superspace. Observe that the reference to 'Euclidean' or 'classically forbidden' regions is rather a matter of convention, since, owing to the indefiniteness of the supermetric, the Euclidean regions are not inaccessible for the ordinary Lorentz paths. It is well known that the Lorentz paths in the cosmological models with scalar field can enter Euclidean regions, and the paths corresponding to the Euclidean equations of motion may pass through the Lorentz region [29, 71, 21]. Nevertheless, the Euclidean and the Lorentz regions make sense in the description of processes of tunneling in cosmology and in the physics of instantons [68] even in case of an indefinite supermetric. The term 'Euclidean region' can even be given a precise meaning if we define it as a region which may host the points of minimum contraction and maximum expansion of the Universe. So, the Euclidean region is a region where

$$U(a,x) > 0. (3.17)$$

Accordingly, the boundary of this region is defined by the equation

$$U(a, x) = 0. (3.18)$$

Instantons are solutions of the Euclidean equations of motion for which both velocities \dot{x} and \dot{a} vanish on the boundary. Such solutions correspond to the extremes of the Euclidean action, and, accordingly, to the extremes in the probability distributions derived from the wave functions.

Now is the time to make a terminological point. Strictly speaking, the concept of a 'no-boundary' wave function cannot be straightforwardly applied to the case of a complex scalar field. The fact is that the prescription 'no-boundary' in the semiclassical approximation reduces to a continual integration from a = 0 to a certain small cosmological radius a with respect to compact metrics and regular matter fields. In the presence of the phase variable θ , however, the associated centrifugal term makes the matter fields not regular at a = 0. Because of this, we need to modify the 'no-boundary' wave function. The modification proposed in Ref. [47] consists in that the wave function of the Universe must grow exponentially in the Euclidean region, and be proportional to $\exp(-I)$ in the semiclassical approximation. The instantons considered in Refs [45-49] correspond to precisely this definition of 'no-boundary' wave function.

Now we can embark on studying the shape of the Euclidean regions in the plane of superspace variables (a, x) for different values of the parameters of the superpotential U (3.11).

In the simplest case $Q = \Lambda = \lambda = \xi = 0$, we get the Euclidean region delimited by the hyperbolic curve $x = \pm \sqrt{3/(4\pi)} m_{\rm P}/(ma)$ (Fig. 1a). Introduction of the cosmological term $\Lambda \neq 0$ closes the Euclidean region on the right at $a = \sqrt{3/\Lambda}$ (Fig. 1b).

Introduction of a non-zero classical charge for the scalar field $Q \neq 0$ closes the Euclidean region on the left, and we get a banana-like shape (Fig. 1c) [45, 46].

After the introduction of a small term accounting for the non-minimal coupling between the scalar field and gravitation $(\xi \neq 0)$, we get another Euclidean region in the upper left-hand corner of the (x, a) plane [48]. This new region is not compact, and is not limited from above (Fig. 1d). As ξ increases, the second Euclidean region drops down, and at a certain value of ξ joins the first banana-shaped Euclidean region (Fig. 1e). As ξ increases further, we get a unified



Figure 1. Geometry of the Euclidean regions for different values of parameters.

Euclidean region whose boundary is partly concave, partly convex (Fig. 1f). Further on, the boundary becomes convex (Fig. 1g). After introduction of the self-interaction constant $\lambda \neq 0$, we may, depending on the values of Q, λ , ξ , and m, get three different configurations of the Euclidean regions. For example, with sufficiently small ξ and sufficiently large λ , we get one banana-shaped Euclidean region (Fig. 1h). With small values of λ , incapable of 'closing' the upper region, we get two separate Euclidean regions (rather similar to Fig. 1d). Finally, with a sufficiently large value of ξ we get an open 'bag' with infinitely long narrow neck. We see that the inclusion of the charge Q, non-minimal coupling ξ , and self-action for the scalar field, offers a great diversity of possible geometries of Euclidean regions.

Now we perform a numerical integration of the Euclidean analogs of Eqns (3.14) and (3.15) to resolve the question concerning the existence of instantons which realize the extremes of Euclidean action. The numerical integration can be supplemented by finding exact expressions for the possible locations of the points of the largest expansion of the Universe ($\dot{a} = 0$, $\ddot{a} < 0$) and the points of its lowest contraction ($\dot{a} = 0$, $\ddot{a} > 0$), as well as the point of maximum and minimum x. Such an analysis has been carried out in detail in Refs [49, 32, 33]. Here it will suffice to state that the instanton paths must certainly cross the curves which separate the points of the largest expansion and the lowest contraction of the Universe, and the points of possible maximum and minimum of x.

It is straightforward that instantons do not exist in the framework of the simplest model with a real scalar field. If, however, the charge Q is non-zero, the instanton may and does exist (Fig. 2a). Using the end point of this instanton on the right-hand boundary of the Euclidean banana-shaped region as the initial point of the Lorentz path, we see that this path exhibits quasi-inflation behavior [45–49].

In the case of a model with a complex scalar field with non-minimal coupling, when there are two separate Euclidean regions (Fig. 1d), the instanton occurring in the bananashaped region may ensure acceptable initial conditions for inflation (Fig. 2b). At the same time, there is another solution of the Euclidean equations of motion with zero velocity on the boundary between the Euclidean and Lorentz regions: this solution connects the Euclidean regions passing through the Lorentz region. The behavior of the Lorentz path which starts at the end point of the Euclidean path is definitely noninflationary, and cannot be used for describing the quantum tunneling of the Universe from nothing.

Further on, as the constant of non-minimal coupling ξ increases, the path made by this unusual instanton becomes shorter. As soon as the two Euclidean regions meet, the second instanton reduces to a point and disappears.

As ξ increases further, when we have a single Euclidean region open from above, there are two instantons inside the Euclidean region (Fig. 2c). The lower instanton corresponds to the local maximum of the magnitude of Euclidean action, and the upper to its local minimum. The end points of both instantons can be used as the starting points of Lorentz paths exhibiting quasi-inflationary behavior (Fig. 2c).

It is important to note that the upper instanton ensures the existence of a peak in the probability distribution of the tunneling wave function of the Universe [18], since in the tree-level approximation this function behaves as $\Psi_{\rm T} \sim \exp(-|I|)$. At the same time, the lower instanton ensures a peak in the probability distribution of the no-boundary wave



Figure 2. Boundaries between Euclidean and Lorentz regions (1); instanton paths (2); Lorentz paths (3); x and a are separating curves.

function [17], which behaves as $\Psi_{\rm NB} \sim \exp(+|I|)$. Thus, if we opt for the no-boundary wave function of the Universe, we must use the end point of the lower instanton for defining the most probable initial conditions of inflation, whereas the upper instanton must be used in the case of the tunneling wave function.

Finally, when ξ is large enough, and the boundary of the Euclidean region is convex (Fig. 1g), there are no instantons at all.

When the interaction term $\lambda \neq 0$ is taken into account, the overall pattern is more or less similar to that described above, with some additional possibilities which are discussed in Refs [49, 32, 33].

It would also be interesting to consider the case of a zero cosmological constant ($\Lambda = 0$). The vanishing of the cosmological constant opens up the Euclidean region on the right (Fig. 2f). In place of two instantons we now have only one, corresponding to the lowest magnitude of Euclidean action and defining the peak of the probability distribution of the tunneling wave function of the Universe. The second instanton turns into a path that continues indefinitely within the Euclidean region, never reaching its boundary. In this case it is only the tunneling wave function of the Universe that can predict the most probable conditions for the start of cosmological evolution, as opposed to the situation with the single Euclidean region, in which preference is given to the noboundary wave function. This fact may add vigor to the discussions between advocates of different versions of the wave function of the Universe [72, 73].

We end this section by noting that the instanton solutions can also be obtained in the framework of the model with a real scalar field by selecting the potential which exhibits extreme points at certain values of the inflaton field. Then we get the instanton solutions with a constant inflaton field and a cosmological factor which varies according to the Euclidean De Sitter law:

$$a(\tau) \sim \sin H \tau$$

where τ is the Euclidean time. On the boundary between the Euclidean and Lorentz regions, these solutions can be analytically extended by accomplishing the analytic transition from Euclidean time to Lorentz time *t*:

$$\tau = \frac{\pi}{2H} + \mathrm{i}t\,,\tag{3.19}$$

$$a(t) \sim \cosh Ht \,. \tag{3.20}$$

The instantons in models with a real scalar field have, however, a considerable flaw: the inflation they predict is eternal (at least when it starts from exactly the end point of the instanton solution). A natural way out of inflation with heating of the Universe and a transition to the Friedmann stage is not available. By contrast, the inflation stage predicted by the model with a complex inflaton field lasts for long but not forever, and allows for an elegant way out of inflation and on to the Friedmann Universe.

4. Energy scale of inflation and loop effects

The existence of instanton solutions in the model described in the previous section allowed a description of the birth of the Universe in a very simple manner. With other models, where such solutions do not exist, the semiclassical analysis of tunneling calls for the introduction of complex time and complex paths. Such complexification becomes necessary even when dealing with many-dimensional tunneling in nonrelativistic quantum mechanics [24]. In the context of the quantum-cosmological problem, if we want to treat the paths as being composed of two (Euclidean and Lorentz) sections — thus introducing the point where Euclidean time converts into Lorentz time — we need that not only the dynamic variables q satisfy the appropriate equations of motion on each section of the path, but also that their derivatives satisfy the following conditions at the point of transition [21-23]:

$$\operatorname{Re} \dot{q}_{\mathrm{E}} = \operatorname{Im} \dot{q}_{\mathrm{L}}; \quad \operatorname{Im} \dot{q}_{\mathrm{E}} = -\operatorname{Re} \dot{q}_{\mathrm{L}}. \tag{4.1}$$

As it turns out, in some models the need for complexification of the dynamic variables does not affect the quantitative aspects of tunneling too much, and its effects can be taken into account using the perturbation theory [22, 23]. In this section we are going to discuss some models of this kind. There are two key aspects:

(1) going beyond the mini-superspace — that is, the inclusion, along with the cosmological radius a and the spatially homogeneous harmonic of the inflaton field φ , of perturbations of these and all other matter fields;

(2) going beyond the tree-level approximation for calculating the wave function of the Universe — that is, the inclusion of loop corrections to the semiclassical formulas (3.2), (3.3).

Such an approach not only allows a description of the quantum birth of the Universe and the prediction of the initial conditions of inflation, but also ensures the normalizability of the wave function of the Universe [22, 23, 38-43], which is a necessary condition for its probabilistic interpretation. In turn, the requirement of normalizability imposes certain restrictions on the spectrum of the theory, which is a non-trivial manifestation of the principle of self-consistency mentioned in the introduction.

Let us now briefly discuss the results of Refs [22, 23, 38-43]. The wave function of the Universe in the one-loop approximation is

$$\Psi(q)_{\mathrm{T,NB}} = \exp\left(\pm I[q] - W[q]\right), \qquad (4.2)$$

where W is the one-loop correction to the effective action:

$$W = \frac{1}{2} \operatorname{Tr}_{\mathcal{M}} \ln \frac{F}{\mu^2}, \qquad F = \frac{\delta^2 I}{\delta \xi \delta \xi}, \qquad (4.3)$$

 $\xi = [\varphi, f], f \ll \varphi$. Here ξ are the physical degrees of freedom, φ is the homogeneous mode of the inflaton field, f are the inhomogeneous modes of all the fields, and μ^2 is the renormalization mass parameter.

The linearized approximation for *I* is

$$I[g] = I + \frac{1}{2} f^{\mathrm{T}}(\mathrm{D}v)v^{-1}f\Big|_{\tau_{+}}.$$
(4.4)

Here $I = I[g_0]$ is a function of the homogeneous mode φ , v is the complete set of the basic functions of operator F, regular on the De Sitter Euclidean sphere (inhomogeneous modes), and D is a differential operator of the first order, Wronskian-related to operator F. It can be demonstrated that the expression for the 'Euclidean' wave function of the Universe is

$$\Psi_{\mathrm{T,NB}}(\tau_{+}|\varphi,f) = \frac{1}{\left[\det u\right]^{1/2}} \exp\left[\pm I - \frac{1}{2}f^{\mathrm{T}}(\mathrm{D}v)v^{-1}f\right].$$
(4.5)

After the analytic extension into Lorentz space-time, the wave function (4.5) becomes

$$\Psi_{\mathrm{T,NB}}(t|\varphi, f) = \left(\frac{1}{\left[\det u^*\right]^{1/2}}\right)^{\mathrm{R}} \\ \times \exp\left[-I_{\mathrm{B}} + \mathrm{i}S + \frac{1}{2}\,\mathrm{i}f^{\mathrm{T}}(\mathrm{D}v)v^{-1}f\right] \quad (4.6)$$

and is a De Sitter invariant vacuum for f modes [74]. The superscript R here denotes the renormalization of the infinite product of basic functions; S is the Lorentz action. It is in this way that the quantum state of matter fields in Lorentz space – time (and Lorentz space – time itself) arises by virtue of quantum tunneling from the classically forbidden Euclidean world.

Information concerning the normalizability of the wave function of the Universe (4.6) can be obtained from the diagonal elements of the density matrix

$$\hat{\rho} = \mathrm{Tr}_f |\Psi\rangle \langle \Psi| \,. \tag{4.7}$$

It has been shown [43] that the quantity (4.7) can be expressed as

$$\rho(t|\varphi) = \frac{\Delta_{\varphi}^{1/2}}{|u_{\varphi}(t)|} \exp(\pm I - \Gamma_{1\text{-loop}}), \qquad (4.8)$$

where $\Gamma_{1-\text{loop}}$ is the one-loop correction to the effective action calculated on the closed compact De Sitter instanton. This quantity is conveniently calculated by the zeta-regularization technique [75], which allows $\Gamma_{1-\text{loop}}$ to be represented as

$$\Gamma_{1-\text{loop}} = -\frac{1}{2}\zeta'(0) - \frac{1}{2}\zeta(0)\ln\mu^2 R_0^2, \qquad (4.9)$$

where $\zeta(s)$ is the generalized Riemann zeta function, and R_0 is the radius of the instanton. The generalized Riemann zeta function is defined as

$$\zeta(s) = \sum \frac{1}{\lambda^s} \,, \tag{4.10}$$

where λ are the eigenvalues of operator *F*. In the limit $\phi \to \infty$ (or, which is the same, $R_0 \to 0$), expression (4.10) reduces to

$$\rho(\varphi) \sim \exp(\pm 2I_{\rm B})\varphi^{-Z-2}, \qquad (4.11)$$

where Z is the anomalous scaling of the theory, expressed in terms of $\zeta(0)$ for all fields included in the model. The requirement of normalizability

$$\int_{-\infty}^{\infty} \mathrm{d}\varphi \,\rho(\varphi) < \infty \tag{4.12}$$

imposes a restriction on Z,

$$Z > -1$$
, (4.13)

which is required for the integral with respect to φ to converge at $\varphi \to \infty$.

The most convenient method for calculating the anomalous scaling is the Schwinger – DeWitt expansion [76], since $\zeta(0)$ is known to coincide with the second coefficient in this expansion:

$$\zeta(0) = A_2 \,. \tag{4.14}$$

A comparison between the different models against the criterion (4.13) on the assumption that the interaction between the inflaton field and all other matter fields can be neglected [40] brings one to the conclusion that supersymmetrical models are more suitable. More interesting, however, is the model with a strong non-minimal coupling between the inflaton field and gravitation [69]. Such a model was considered in Refs [22, 41, 42]. The Lagrangian of this model is similar to that given by Eqn (3.4), but the inflaton field is real. In addition, the matter fields interact with the inflaton field, and in doing so acquire large masses at the initial (inflation) stage of cosmological evolution. This model not only ensures the normalizability of the wave function of the Universe, but also features a peak in the probability distribution corresponding to this wave function. It turns out that the no-boundary wave function of the Universe displays a peak in the probability distribution only with a specially selected potential, and then inflation is eternal there is no elegant way out. By contrast, the tunneling wave function exhibits a probability peak at a certain value of the inflaton field φ_{I} , and the required duration of the inflation stage (cosmological observations indicate that the Universe during the inflation stage ought to expand by a factor of e^{60}) is ensured if the anomalous scaling of the model has a huge value of

$$Z \sim 10^{11}$$
 (4.15)

This value can be obtained naturally enough in the model under consideration, where the leading contribution to Z is

$$Z = 6 \frac{\xi^2}{\lambda} A + O(\xi), \qquad (4.16)$$

where

$$A = \frac{1}{2\lambda} \left(\sum_{\chi} \lambda_{\chi}^{2} + 4 \sum_{A} g_{A}^{4} - 4 \sum_{\psi} f_{\psi}^{4} \right).$$
(4.17)

Here λ is the constant of self-action of the inflaton field; λ_{χ} , g_A and f_{ψ} are the constants of interaction of the inflaton with scalar, vector and spinor fields, respectively. Since the usual values for parameters for models with non-minimal coupling [69] are $\lambda \simeq 0.05$, $\xi \simeq -2 \times 10^4$, $\lambda/\xi^2 \simeq 10^{-10}$, the required value of Z is obtained if the parameter A, which depends on the selected model of the physics of elementary particles, is fixed within rather stringent limits:

This restriction once again points to a quasi-supersymmetric nature for the model of the physics of elementary particles, which has to maintain a rather precise balance between bosons and fermions [see Eqn (4.17)].

Further and much more extensive analysis was carried out in Ref. [23], where not only the anomalous scaling was calculated, but also the contribution from the function $\zeta'(0)$, with due account for the distortion of the shape of the instanton (caused by the time dependence of the inflaton field). The main results of the earlier works [22, 41, 42] were not much modified. Observe, however, that when speaking of quasi-supersymmetry we only mean some balance between bosons and fermions rather than a final selection of any particular model of supersymmetry. Moreover, in Ref. [43] it was demonstrated that it is hard to satisfy the criterion of normalizability (4.13) in models with exact supersymmetry.

To end this section, we note that the results of Refs [22, 23, 38-43] point to the existence of a deep relationship between quantum gravitation, inflation cosmology, and the particle physics — combining the requirements which follow from the general principles of quantum theory with observations interpreted in the light of quantum cosmology, one may draw certain conclusions regarding the structure of acceptable models of the particle physics.

5. Hamiltonian quantization of gravitation and quantum self-consistency

In a sense, quantum cosmology occupies a unique place in contemporary theoretical physics. On the one hand, it is linked (mainly through the inflation models [61, 37]) with observational cosmology, including such successful branches as the study of the microwave background and the large-scale structure of the Universe. On the other hand, the mathematical structure of quantum cosmology puts it close to such branches of theoretical physics as the theory of superstrings [35], the theory of membranes, etc. This last aspect seems to us somewhat underexploited, and in this section we are going to give a brief account of one attempt [44] to apply the new methods tried out in the theory of strings to quantum cosmology.

We are referring to the study of quantum anomalies using the Hamiltonian BFV-BRST (Batalin-Fradkin-Vilkovisky-Becchie-Rue-Stora-Tyutin) quantization [77], which has proved its efficiency in the theory of strings [78]. We ought to add that for us the string theory was a well of experience rather than a source of particular forms of Lagrangians, as opposed to the currently popular string cosmology [57]. To wit, we borrowed the idea concerning the possible opening of the quantum algebra of constraints from the string theory, as well as the existence of the critical relations between the parameters of the theory. This idea was applied to Einstein's theory of gravitation.

Recall that the main equations of canonical quantum gravitation and cosmology include, in addition to the Wheeler – DeWitt equation (3.1), the so-called equations of supermomenta

$$H_{\rm i}|\psi\rangle = 0\,,\tag{5.1}$$

which correspond to the invariance of the wave function of the Universe with respect to spatial diffeomorphisms. It is usual to require that Eqns (3.1) and (5.1) be satisfied simultaneously. This approach goes back to Dirac's quantization of systems with constraints. As a matter of fact, these constraints at the classical level are constraints of the first kind, and are involutive with respect to the Poisson brackets [80]. When, however, we consider operators and their commutators, the algebra may become open, and Dirac's quantization is not feasible. Such a situation is well known in the theory of strings, where (for example, in the quantum theory of boson string) the closed algebra of constraints of the first kind is replaced with a centrally extended Virasoro algebra. Accordingly, only half of the constraints eliminate the physical states, and the quantum theory is only selfconsistent when the dimensionality of the embedding space is

$$D = 26.$$
 (5.2)

One way to obtain Eqn (5.2) is to use the Hamiltonian BRST quantization [77]. In the context of this approach, the verification of self-consistency reduces to some straightforward algebra.

An attempt at Hamiltonian BFV–BRST quantization of closed cosmological models in Ref. [44] took advantage of the similar symmetries of the theory of strings and the general relativity (both theories are reparametrization-invariant). The proposed method reduces to the following scheme:

(1) both the dynamic variables of the theory and the constraints are expanded in harmonics;

(2) subalgebra of area-conserving diffeomorphisms is extracted from the algebra of constraints, and the remaining constraints are used for constructing the so-called Virasorolike generators;

(3) the model is selected: the calculations turn out to be simplest in the so-called extended Bianchi I stationary model, in which space is represented as a time-independent *N*-dimensional torus;

(4) the constraints and structure functions of the pseudoalgebra of constraints are expanded in perturbation theory in the small parameter $l_P/V^{1/N}$, where l_P is the Planck length, and V is the volume of the Universe;

(5) the BRST operator Ω is constructed and the quantum condition of nilpotency $\hat{\Omega}^2 = 0$ verified;

(6) this verification on the one-loop level (that is, the calculation of the first quantum correction to commutators of constraints) results in the following critical relation:

$$d = 30 + \frac{5}{2}(N+1)(N-2).$$
(5.3)

Here d is the number of massless scalar fields in the theory. The generalization of Eqn (5.3) for a field of arbitrary spins will be given below.

Let us briefly comment on the steps described above. Considering compact manifolds, one may express phase variables and constraints in terms of a discrete set of coefficients of expansion in harmonics. The functional base of this expansion is formed by the eigenfunctions of the Laplace operator defined over the maximally symmetrical space of the given topology (in the case of the extended Bianchi I model, these functions are multi-dimensional Fourier harmonics on a torus). While the expansion of dynamic variables in harmonics is standard in cosmology [81, 10], the expansion of constraints with respect to a discrete base is a novel feature. Contracting the generators of supermomentum (5.1) with the transverse vector harmonics, we obtain the generators of area-conserving diffeomorphisms introduced by V I Arnold in hydrodynamics [82], and widely employed in the p-brane theories [83]. The remaining supermomenta (longitudinal) together with the super-Hamiltonian can be used for constructing the Virasoro-like generators, similar to the Virasoro generators in the theory of the closed boson string. The structure coefficients in the new discrete base are expressed in terms of the integrals of triple products of the relevant harmonics, which in turn are expressed in terms of the quadratic combinations of Clebsch-Gordan coefficients of the relevant symmetry group in accordance with the Wigner-Eckart theorem [84]. These coefficients are very easy to calculate for the Bianchi I model. Observe that, apart from its computational expedience, a torus topology for the Universe is physically attractive for cosmologists, and quite a few papers have been concerned with the observational limitations on the feasibility of such a model [85].

The expansion of the metric and the conjugate momentum with respect to the operators of creation and annihilation, and the subsequent expansion of the constraints in powers of these operators, naturally give rise to a small parameter $l_{\rm P}/V^{1/N}$. Amazingly, the higher-order corrections to commutators of Virasoro-type constraints are multiplied by this parameter, thus justifying the use of the perturbation theory.

All information concerning the gauge symmetry of the system can be formulated in the BFV-BRST formalism. In the classical theory, one can construct a Grassmann object — the BRST charge

$$\Omega = c^{\alpha} T_{\alpha} + \frac{1}{2} P_{\gamma} U^{\gamma}_{\alpha\beta} c^{\beta} c^{\alpha} , \qquad (5.4)$$

where T_{α} are constraints of the first kind, $U_{\alpha\beta}^{\gamma}$ are the structural constants of the algebra of constraints, c^{α} are the Faddeev–Popov ghosts, and P_{α} are the conjugate momenta. The disappearance of the generalized Poisson bracket or the classical nilpotency of the BRST charge are equivalent to the involution relations between the constraints

$$\{T_{\alpha}, T_{\beta}\} = U^{\gamma}_{\alpha\beta}T_{\gamma},$$

and the Jacobi identities.

Going over to quantum operators, one may impose the following condition on the quantum states:

$$\hat{\Omega}|\psi\rangle = 0$$
,

which is less stringent than Dirac's annihilation of the physical state by all constraints. The condition of quantum nilpotency or quantum self-consistency

$$\hat{\Omega}^2 = 0$$

carries information on quantum anomalies and conditions for their cancellation owing to the ghosts' contribution. It is from this condition that one can get the critical relations, like d = 26 for boson strings, or Eqn (5.3).

For the model in question, the BRST charge involves about a hundred terms. Ghosts corresponding to the Virasoro-like generators are normally ordered (according to Wick), whereas ghosts corresponding to area-conserving diffeomorphisms are Weil-quantizable. It is this choice that is self-consistent. The analysis of nilpotency condition $\hat{\Omega}^2 = 0$ leads to the following condition of quantum self-consistency:

$$d + d_{\rm V}(N-1) + d_{\rm F} 2^{(N-1)/2-1} = 30 + \frac{5}{2}(N-2)(N+1).$$
(5.5)

Here *d* is the number of massless scalar fields, d_V is the number of massless fields of spin 1, and d_F is the number of massless fermion fields. Analyzing the implications of this result, we would like to draw attention to the following:

1. Formula (5.5) is much simpler than the method used to derive it, and has a rather natural structure. Observe that (1/2)(N-2)(N+1) is the number of transverse-traceless gravitons in the (N+1)-dimensional space-time. It is important that the coefficients on the right-hand side of Eqn (5.5) are determined by the selection of vacuum for gravitons.

2. Formula (5.5) has a reasonable limit at N = 1 (the boson string). If N = 1, then d = 25. This result can be easily

understood from the standpoint of the string σ -model without Weil invariance [86].

3. For any N, an empty steady closed Universe of the Bianchi I type is not quantum self-consistent.

4. With N = 3 (our Universe) the number of degrees of freedom of matter fields is 40. This is obviously too small even for the Standard Model, to say nothing of the models of the Grand Unification or supersymmetrical models. This, however, should not discourage us, since we are almost certain that our Universe is not toroidal, and in any case it is not steady.

Let us finish this section with a list of possible directions of further development of this formalism:

(1) massive fields (which is a non-trivial task from a technical standpoint);

(2) a non-stationary (expanding) Universe;

(3) a Universe with different topologies (in the first place, the Bianchi IX model);

(4) Kaluza – Klein models.

6. Dynamics of cosmological models with a scalar field: chaos and fractality

In this section we shall consider a simple model whose action can be found from Eqn (3.4), if we replace the complex scalar field with a real scalar field and set $\xi = \lambda = \Lambda = 0$. The equations and the first integral of motion for this system are

$$\frac{m_{\rm P}^2}{16\pi} \left(\ddot{a} + \frac{\dot{a}^2}{2a} + \frac{1}{2a} \right) + \frac{a\dot{\varphi}^2}{8} - \frac{m^2\varphi^2 a}{8} = 0 \,, \tag{6.1}$$

$$\ddot{\varphi} + \frac{3\dot{\varphi}\dot{a}}{a} + m^2\varphi = 0, \qquad (6.2)$$

$$-\frac{3}{8\pi}m_{\rm P}^2(\dot{a}^2+1)+\frac{a^2}{2}(\dot{\varphi}^2+m^2\varphi^2)=0.$$
(6.3)

Looking at Eqn (6.3), we see at once that the points of maximal expansion and minimal contraction may only exist in the Euclidean region, where

$$\varphi^2 \leqslant \frac{3}{4\pi} \frac{m_{\rm P}^2}{m^2 a^2} \,, \tag{6.4}$$

which is a region in the half-plane $0 \le a < +\infty$, $-\infty < \phi < +\infty$ delimited by the hyperbolic curves $\phi \le \sqrt{3/4\pi} m_{\rm P}/m_a$ and $\phi \ge -\sqrt{3/4\pi} m_{\rm P}/m_a$ (Fig. 3). Now we have to find the possible locations of the points of minimum contraction ($\dot{a} = 0$, $\ddot{a} > 0$), and maximum expansion ($\dot{a} = 0$, $\ddot{a} < 0$). If $\dot{a} = 0$, then $\dot{\phi}^2$ can be expressed from Eqn (6.3). Substituting the resulting value of $\dot{\phi}^2$ and $\dot{a} = 0$ into Eqn (6.1), we get

$$\ddot{a} = \frac{4\pi m^2 \varphi^2 a}{m_{\rm P}^2} - \frac{2}{a} \,. \tag{6.5}$$

From this equation we see that the possible points of maximum expansion lie within the region

$$\varphi^2 \leqslant \frac{1}{2\pi} \frac{m_{\rm P}^2}{m^2 a^2},$$
(6.6)

whereas the possible points of minimum contraction lie outside this region but not beyond the Euclidean region (6.4) (see Fig. 3).



Figure 3. Solid hyperbolic curves separate Euclidean and Lorentz regions in the half-plane $a \ge 0$, φ . Dashed hyperbolic curves separate the possible points of maximum expansion from the possible points of minimum contraction.

It is important that all possible paths satisfying Eqns (6.1)-(6.3) contain the points of maximum expansion [87], which allows us to use these points for constructing a convenient classification scheme. The locations of points of maximum expansion as defined by Eqn (6.5) have been earlier described in the language of phase space in Ref. [71], and in the language of configuration space in Ref. [29]. General expressions describing these points for a broad class of models can be found in Ref. [49].

We shall distinguish between the paths which are monotonic with respect to a and φ (that is, falling into the singularity with a monotonically varying magnitude of the scalar field), the bouncing paths (passing through the points of maximum contraction), and those displaying extremes with respect to φ (we shall call them φ -turns). Observe that there is a certain pattern in the arrangement of the points of maximum expansion corresponding to different types of paths. Namely, the regions corresponding to the paths which fall into the singularity after a number of φ -turns, alternate with those corresponding to the bouncing paths.

We are going to analyze the paths starting at certain points of maximum expansion without looking into their history. For velocities $\dot{\phi}$ at a point of maximum expansion we select the 'upward' direction — that is, the sign is positive (the pattern with the downward velocities $\dot{\phi}$ can be obtained by mirror reflection with respect to $\varphi = 0$). Numerical integration of the equations of motion reveals that for the points of maximum expansion located in region θ in Fig. 4 the paths steadily approach the singularity a = 0, $\varphi = +\infty$ (some typical paths of this kind are shown in Fig. 5a). This appears quite natural, since in the region close to the axis of ordinates a = 0 the system behaves like the model with a massless scalar field, where bounces and φ -turns are not possible. The paths that start in region 1 in Fig. 4 exhibit a bounce. Typical paths of this kind are shown in Fig. 5b. Then we have region $\overline{1}$ not shown in the diagram; there is a bounce preceded by a φ -turn. Typical paths of this kind are shown in Fig. 5d. The boundary between regions 1 and $\overline{1}$ is a locus where some periodical paths which evade singularity may pass through the points of



Figure 4. Structure of the region of localization of the possible points of maximum expansion.



Figure 5. Paths starting at the points of maximum expansion in the regions shown in Fig. 4.

maximum expansion. (Such paths were first discovered in Ref. [28] and analyzed in detail in Ref. [29].) In particular, the point $\varphi = 0$ on the boundary between regions I and \overline{I} is crossed by a periodical path symmetrical with respect to φ , shown in Fig. 5c.

Region 1 hosts those paths which do not exhibit a bounce, but fall into the singularity a = 0, $\varphi = -\infty$ after a φ -turn (Fig. 5e).

Now in region 2 (see Fig. 4) we have points of maximum expansion which belong to the paths that, after a φ -turn, bounce off at $\varphi < 0$, thus avoiding the singularity (Fig. 5f). In region 2' we have paths that bounce immediately after two φ -turns. The boundary between regions 2 and 2' contains paths that involve turnpoints, where we simultaneously have $\dot{a} = 0$, $\dot{\varphi} = 0$. Some of these paths are periodical; one is shown in Fig. 5g.

Starting at the points of maximum expansion in region 2' (see Fig. 4), we have paths without bounces but with two φ -turns, which eventually fall into the singularity a = 0, $\varphi = +\infty$ (Fig. 5h).

Now it is easy to understand what the paths are corresponding to the points of maximum expansion in regions 3, $\overline{3}$, 3', 4, $\overline{4}$, 4', etc.

Observe that the most peculiar is the boundary between region I and other regions in the upper half-plane. This boundary is composed of two curves (see Fig. 4). The left-hand curve is finite and separates region I from region I', whereas the right-hand curve is infinite and lies almost parallel to the hyperbolic curve separating the possible points of maximum expansion from the points of minimum contraction. Observe also that this upper curve touches upon all regions starting with region I'.

Figure 5i shows a path whose point of maximum expansion occurs on the upper branch of region *1*. This path exhibits a φ -turn right after a bounce, passes through the second point of maximum expansion, and falls into the singularity in the lower half-plane.

Recall that our classification is based on the 'upward' direction. The structure of regions corresponding to the 'downward' direction can be obtained by a mirror reflection with respect to $\varphi = 0$. We shall mark the 'upward' and 'downward' directions with arrows \uparrow and \downarrow .

The most interesting are the bouncing paths, especially those which feature many (ideally, infinitely many) bounces — that is, the paths that avoid falling into the singularity.

Let us now pay attention to those paths whose points of maximum expansion are located in regions $1, 2, 3, \ldots$, or $1', 2', 3', \ldots$ — that is, the paths that experience at least one bounce. We look at the structure of these regions in the light of where the second point of maximum expansion is located. The substructure of region $1\uparrow$ is shown in Fig. 6, and the shapes of the paths corresponding to different subregions of regions 1 and 1' are shown in Fig. 7.

To the right of the boundary with region 1', we have a subregion $1\uparrow 1'\downarrow$ which corresponds to the paths that after a



Figure 6. Structure of subregions of region 1. Narrow regions denoted b, c, d, e correspond to the paths which feature at least two bounces. The subregion a and other subregions between the narrow ones correspond to the paths that fall into the singularity after one bounce.



Figure 7. Paths corresponding to different subregions of region *I* shown in Fig. 6.

bounce come to the point of maximum expansion in region $1'\downarrow$, and experience another bounce. The next subregion to the left of subregion $1\uparrow 1'\downarrow$ is $1\uparrow 1'\downarrow$, which corresponds to the paths that after a bounce come to the point of maximum expansion in region $1'\downarrow$. Passing through this point of maximum expansion, they feature a φ -turn and arrive at the singularity $a = 0, \varphi = -\infty$.

Next goes subregion $1 \uparrow 1 \downarrow$, corresponding to paths with two bounces. Then follows subregion $1 \uparrow 1' \uparrow$, and then $1 \uparrow 2 \uparrow$, $1 \uparrow 2' \uparrow$, $1 \downarrow 2' \uparrow$, $1 \uparrow 1 \uparrow$, $1 \uparrow 2' \downarrow$, $1 \uparrow 3 \downarrow$, etc.

The structure of region $1'\uparrow$ is much simpler: we only have two subregions, $1'\uparrow 1\downarrow$ and $1' 0\downarrow$.

We see that regions $1\uparrow$, $1'\uparrow$ contain infinitely many subregions, whose structure is more or less similar to the structure of the entire region of possible points of maximum expansion. This structure is, however, kind of 'triplicated'. Indeed, the number of regions is doubled owing to the signs \uparrow and \downarrow . On top of that, we have an infinite number of intermediate regions 11. It can be demonstrated that the structure of pairs of regions 22', 33', 44',... is quite similar to the structure of the pair 11' as described above.

Analyzing the structure of subregions corresponding to two bounces, we find that each of them has a 'subsubstructure' of infinitely many 'sub-subregions'. Proceeding *ad infinitum*, we find that the region of localization of points of maximum expansion, corresponding to the paths that avoid falling into singularity, is a result of an infinite recurrent process, each stage of which generates self-similar structures. Such self-similarity of structures arising at different scales points to the fractal nature of the set obtained through this infinite process [88]. In this way, although the set of paths that avoid singularity and wander eternally between the points of minimum contraction and maximum expansion has measure zero on the set of all paths, it may at the same time have a non-trivial fractal dimension. This phenomenon was first discussed in somewhat different language in Ref. [29].

The analysis of observations indicates that there may currently exist a small cosmological constant [89]. In this connection it would be interesting to analyze the dynamics of the cosmological model which, along with the scalar field, may feature a cosmological constant Λ — or, which is the same, a constant term in the potential of the scalar field. Such an analysis was carried out in Ref. [31].

First of all, we ought to observe that in this case the conditions of the theorem stating the impossibility of infinite expansion of the Universe are violated, and the phase space features two singular points (attracting and repelling focuses), corresponding to expanding and contracting De Sitter Universes. If the field is massless, the picture becomes especially clear: the focuses become nodes, and there are two stationary saddle points, where the radius of the Universe is fixed, and the magnitude of the scalar field grows linearly. This picture corresponds to the known phase diagram for the cosmological model with an ideal liquid and cosmological constant [90], which is not surprising since the equation of state for a massless scalar field is $p = \epsilon$, as indicated in Section 2 of this review.

In this case, we have three classes of paths:

(1) paths coming out of the repelling De Sitter point (that is, paths that start to contract exponentially at very large values of the cosmological radius), which pass through the point of minimum contraction and go towards the attracting De Sitter point (that is, to the regime of exponential expansion);

(2) paths coming out of the repelling De Sitter point and falling into the singularity, and paths coming out of the singularity and expanding to the attracting De Sitter point;

(3) paths that start and end at a singularity.

If we include the mass of the scalar field, the picture becomes more complicated since, like in the case without the cosmological constant, the paths may wander between singularities. Let us comment on the main features of this picture. The shape of the boundary between Euclidean and Lorentz regions is changed, and is defined by

$$\varphi^2 = \frac{3m_{\rm P}^2}{4a^2} - \frac{\Lambda m_{\rm P}^2}{4m^2} \,. \tag{6.7}$$

The Euclidean region occurs exclusively to the right for

$$a = \sqrt{\frac{3}{\Lambda}}$$

The region of localization of the possible points of maximum expansion is separated from the possible points of minimum contraction by the curve

$$\varphi^2 = \frac{m_{\rm P}^2}{2\pi m^2 a^2} - \frac{\Lambda m_{\rm P}^2}{4\pi m^2} \,, \tag{6.8}$$

which crosses the abscissa at

$$a = \sqrt{\frac{2}{\Lambda}}$$

When the cosmological constant Λ is smaller than the mass of the scalar field m, the structure of regions containing the points of maximum expansion corresponding to different regimes resembles the structure described above. Namely, there is an infinite number of regions which host an infinite number of subregions, and so forth. It is interesting that an infinite number of regions 'fit' into the region limited from the right, the right-hand boundary being a certain periodical path that behaves as a repeller [31]. Like in the case without the cosmological constant, the set of paths that avoid falling into singularity has a fractal nature.

However, at a certain large $\Lambda \approx m^2$ the picture is changed dramatically: the structure with an infinite number of subregions disappears, and only two regions remain: region number θ , from which the paths go into singularity, and region number I. The paths whose points of maximum expansion occur in region I experience a bounce and go to the De Sitter attracting focus. This means that with large values of Λ the set of paths that avoid falling into singularity no longer has a fractal nature.

7. Perturbation methods in cosmology

In this section we discuss an attempt to construct a perturbation theory with respect to the exact solution of the set of equations simplified for describing certain qualitative effects presented in the preceding section. Namely, we take advantage of the fact that it is possible to obtain exact solutions for the closed Friedmann model with a spatially homogeneous massless scalar field.

It is convenient to use the conformal time

$$\eta = \int \mathrm{d}t \, \frac{1}{a(t)} \, .$$

The equations of motion (6.2), (6.3) are rewritten in the form

$$\varphi'' + \frac{2\varphi'a'}{a} + m^2\varphi a^2 = 0$$
(7.1)

and

$$\frac{a'^2}{a^2} = -1 + \frac{4\pi}{3m_{\rm P}^2} \,\varphi'^2 + \frac{4\pi}{3m_{\rm P}^2} \,m^2 \varphi^2 \,. \tag{7.2}$$

Here the prime '' denotes differentiation with respect to η . With m = 0, Eqns (7.1) and (7.2) become

$$\varphi'' + \frac{2\varphi'a'}{a} = 0 \tag{7.3}$$

and

$$\frac{a'^2}{a^2} = -1 + \frac{4\pi}{3m_{\rm P}^2} \,\varphi'^2\,. \tag{7.4}$$

This set of equations can be integrated in quadratures. Like in the previous section, for the starting point of evolution we select the time when the Universe is at the point of maximum expansion. Then we get

$$a = a_0 \sqrt{\cos 2\eta} \,, \tag{7.5}$$

$$\frac{a'}{a} = -\tan 2\eta \,, \tag{7.6}$$

$$\varphi' = \sqrt{\frac{3m_{\rm P}^2}{4\pi}} \frac{1}{\cos 2\eta} \,, \tag{7.7}$$

and finally

$$\varphi = \varphi_0 + \frac{1}{4} \sqrt{\frac{3m_{\rm P}^2}{4\pi}} \ln \frac{1 + \sin 2\eta}{1 - \sin 2\eta} \,. \tag{7.8}$$

These solutions correspond to those paths that start at one of the singularities a = 0, $\varphi = \pm \infty$, pass through the point of maximum expansion $a = a_0$, $\varphi = \varphi_0$, and fall into the other singularity a = 0, $\varphi = \mp \infty$.

In the language of the previous section we may say that the entire half-plane (a, φ) is occupied solely by region θ .

It is interesting that, upon transition to Euclidean time, these solutions describe the well known wormhole of S B Giddings and A Strominger [91], while a_0 becomes the radius of the entrance of the wormhole.

Now we construct the perturbation theory for solutions of Eqns (7.1), (7.2), using the solutions of the massless equations (7.5)–(7.7) for the zero approximation. To simplify the calculations, we confine ourselves to the symmetrical case of $\varphi(0) = 0$. In place of variable *a* we use variable *h* defined as

$$h \equiv \frac{a'}{a} \,. \tag{7.9}$$

Representing the solutions of equations of motion as

$$\rho = \varphi^{(0)} + \delta\varphi \tag{7.10}$$

and

$$h = h^{(0)} + \delta h$$
, (7.11)

where $\varphi^{(0)}$ and $h^{(0)}$ are given by Eqns (7.8) and (7.6), respectively, we get the following equations in $\delta\varphi$ and δh :

$$\delta\varphi'' + 2\delta\varphi' h^{(0)} + 2\delta h\varphi^{(0)\prime} + m^2 a^{(0)2}\varphi^{(0)} = 0$$
(7.12)

and

$$h^{(0)}\delta h = \frac{4\pi}{3m_{\rm P}^2} \,\varphi^{(0)} \,\delta \phi' + \frac{4\pi}{3m_{\rm P}^2} \frac{m^2 \varphi^{(0)2} a^{(0)2}}{2} \,. \tag{7.13}$$

Substituting the explicit expressions for $\varphi^{(0)}$, $\varphi^{(0)'}$, $h^{(0)}$, $a^{(0)}$, from Eqns (7.5)–(7.8) into these equations, we get

$$\delta \varphi'' - 2 \tan 2\eta \delta \varphi' + 2\delta h \sqrt{\frac{3m_{\rm P}^2}{4\pi}} \frac{1}{\cos 2\eta} + \frac{1}{4} m^2 a_0^2 \cos 2\eta \sqrt{\frac{3m_{\rm P}^2}{4\pi}} \ln \frac{1 + \sin 2\eta}{1 - \sin 2\eta} = 0$$
(7.14)

and

$$h^{(0)}\delta h = \sqrt{\frac{4\pi}{3m_{\rm P}^2}} \frac{1}{\cos 2\eta} \,\delta\varphi' + \frac{m^2 a_0^2}{32} \ln^2 \frac{1+\sin 2\eta}{1-\sin 2\eta} \,. \tag{7.15}$$

From these equations we see that the small parameter of the perturbation theory is $m^2 a_0^2$. Using Eqn (7.15), we express δh via $\delta \varphi$:

$$\delta h = -\sqrt{\frac{4\pi}{3m_{\rm P}^2}} \frac{\delta \varphi'}{\sin 2\eta} - \frac{m^2 a_0^2}{32} \frac{\cos^2 2\eta}{\sin 2\eta} \ln^2 \frac{1 + \sin 2\eta}{1 - \sin 2\eta} \,. \tag{7.16}$$

Substituting δh from Eqn (7.16) into Eqn (7.14), we arrive at the following equation in $\delta \varphi$:

$$\delta\varphi'' - 2\delta\varphi' \left(\tan 2\eta + \frac{1}{\cos 2\eta \sin 2\eta} \right) - \sqrt{\frac{3m_{\rm P}^2}{4\pi}} m^2 a_0^2 \left(\frac{\cos 2\eta}{16\sin 2\eta} \ln^2 \frac{1 + \sin 2\eta}{1 - \sin 2\eta} - \frac{\cos 2\eta}{4} \ln \frac{1 + \sin 2\eta}{1 - \sin 2\eta} \right) = 0.$$
(7.17)

The solution of this equation is

$$\begin{split} \delta\varphi' &= \sqrt{\frac{3m_{\rm P}^2}{4\pi}} \frac{m^2 a_0^2}{16} \frac{\sin 2\eta}{\cos^2 2\eta} \\ &\times \left[2\sin 2\eta - \frac{1 + \sin^2 2\eta}{2\sin 2\eta} \ln^2 \frac{1 + \sin 2\eta}{1 - \sin 2\eta} \right. \\ &- \cos^2 2\eta \ln \frac{1 + \sin 2\eta}{1 - \sin 2\eta} + \ln^2 (1 - \sin 2\eta) \\ &- \ln^2 (1 + \sin 2\eta) + 2\ln 2\ln \frac{1 + \sin 2\eta}{1 - \sin 2\eta} \\ &+ 2\text{Li}_2 \left(\frac{1 - \sin 2\eta}{2} \right) - 2\text{Li}_2 \left(\frac{1 + \sin 2\eta}{2} \right) \right], \end{split}$$
(7.18)

where $Li_2(x)$ is the dilogarithm represented as a series

$$\mathrm{Li}_2(x) \equiv \sum_{n=1}^{\infty} \frac{x^n}{n^2} \,.$$

Analyzing the behavior of δh and $\delta \varphi'$, we see [30] that both first corrections have signs opposite to the signs of the respective functions in the zero approximation. We may say therefore that the inclusion of the first corrections $\delta \varphi'$ and δh will describe a transition from the set of paths without bounces and φ -turns to the more sophisticated paths. It can be demonstrated that the absolute value of $\delta h/h^{(0)}$ is everywhere greater than $\delta \varphi'/\varphi^{(0)}$. This means that the bounce occurs before the φ -turn — accordingly, our perturbation theory describes the transition from region θ to region I (see Section 6). (Obviously, the first approximation of the perturbation theory only describes a few of the total number of paths, since the regions located to the right of region Icorrespond to large values of parameter $m^2 a_0^2$, and cannot be described by the perturbation theory.)

Admittedly, this perturbation scheme stops working when the singularity is approached. Indeed, $\delta h/h^{(0)}$ reaches the value of (-1) at *any* value of ma_0 , which means that any path features a bounce irrespective of the initial conditions. We know, however, that for small a_0 the paths lie in region θ , and thus there are neither bounces nor φ -turns. Unfortunately, we are not able to evaluate the parameters which characterize the switch from one regime to another, since our perturbation theory does not work in the neighborhood of singularity. Nevertheless, we may draw some reasonable conclusions from the combination of perturbation methods, numerical calculations, and equations which describe the locations of extreme points of the paths [30].

We end this section by noting that the perturbation theory proposed here for the simplest model with a real scalar field is much less efficient than the perturbation theory which describes the oscillatory approach to singularity (see Section 2). This is because the perturbations $\delta \varphi$ and δh grow in such a way that at a certain time η they become greater than the solutions themselves in the massless approximation. The perturbation theory, so to say, destroys itself. This situation can be compared to the known problem of null-charge in quantum electrodynamics, where the one-loop correction to the constant of electromagnetic interaction is greater than its initial 'tree-level' value [93]. Although this analogy is as remote as the analogy between the application of perturbation theory to the oscillatory approach to singularity and the phenomenon of asymptotic freedom, mentioned in Section 2, it is interesting that these two types of perturbation theory — self-supporting and self-destructing — are found in different branches of classical and quantum physics.

8. Conclusions

We have discussed such problems of quantum and classical cosmology as the oscillatory approach to singularity, the construction of the wave function of the Universe by analyzing the process of quantum tunneling, the study of probability distributions for the initial conditions of cosmological evolution, the calculation of quantum corrections on cosmology, the analysis of quantum self-consistency of different cosmological models, the analysis of classical dynamics of isotropic cosmological models, and the construction of non-singular cosmological scenarios. The status of these problems is not the same: while the study of the oscillatory approach to singularity has become a separate chapter of mathematical physics, the analysis of quantum self-consistency of different cosmological models has only just begun. In general, however, we believe that both quantum and classical cosmologies are the most dynamic branches of modern theoretical physics. We have tried to trace the modern developments in cosmology back to the ideas which Landau put forward at different times. We would also like to observe that the study of chaos, actively pursued now in different branches of physics, including such cosmological aspects as the dynamics of isotropic and anisotropic models, was also one of the topics of interest to Landau, who made an important contribution to the study of a major problem: the problem of turbulence in hydrodynamics [94].

Acknowledgments

The authors are grateful to A O Barvinsky, A A Starobinsky, A V Toporensky for fruitful discussions. This work was supported by the Russian Foundation for Basic Research (RFBR) through Grants 96-02-16220 and 96-02-17591, Grant 96-15-96458 for the support of leading research schools, and Grant RFBR–INTAS 95-1353. One of the authors (I Khalatnikov) wishes to thank the School of Physics and Astronomy of the University of Tel Aviv for financial support.

References

- 1. Landau L D, Private communication
- 2. Bardeen J, Cooper L N, Schrieffer J R Phys. Rev. 108 1175 (1957)
- Wilson K G, Kogut J "The renormalization group and the εexpansion" *Phys. Rep.* 12 75 (1974) [Translated into Russian (Moscow: Mir, 1975)]
- Penrose R Structure of Space Time (New York, Amsterdam: W.A. Benjamin, 1968) [Translated into Russian (Moscow: Mir, 1972)]; Hawking S W, Ellis G F R The Large Scale Structure of Space – Time (Cambridge, New York: Cambridge Univ. Press, 1973)

5.

[Translated into Russian (Moscow: Mir, 1977)]; Hawking S W, Penrose R Proc. R. Soc. London Ser. A **314** 529 (1970)

- Khalatnikov I M, Lifshitz E M Phys. Rev. Lett. 24 76 (1970)
- Lifshitz E M, Khalatnikov I M Pis'ma Zh. Eksp. Teor. Fiz. 11 200 (1970) [JETP Lett. 11 123 (1970)]
- Belinskii V A, Lifshitz E M, Khalatnikov I M Usp. Fiz. Nauk 102 463 (1970) [Sov. Phys. Usp. 13 745 (1971)]
- Belinsky V A, Khalatnikov I M Zh. Eksp. Teor. Fiz. 56 1700 (1969) [Sov. Phys. JETP 29 911 (1969)]
- 9. Landau L D, Lifshitz E M *Teoriya Polya* (The Classical Theory of Fields) (Moscow: Nauka, 1988) [Translated into English (Oxford: Pergamon Press, 1975)]
- Lifshitz E M, Khalatnikov I M Usp. Fiz. Nauk 80 391 (1963) [Sov. Phys. Usp. 6 495 (1964)]
- 11. Lifshitz E M, Lifshitz I M, Khalatnikov I M Zh. Eksp. Teor. Fiz. **59** 322 (1970) [Sov. Phys. JETP **32** 173 (1971)]
- Belinskii V A, Khalatnikov I M, Lifshitz E M Adv. Phys. 31 639 (1982)
- Khalatnikov I M, Lifshitz E M, Khanin K M, Shchur L N, Sinai Ya G J. Stat. Phys. 38 97 (1985)
- 14. Misner C W Phys. Rev. Lett. 22 1071 (1969)
- 15. Cornish N J, Levin J J Phys. Rev. D 55 7489 (1997)
- Cotsakis S, Lemmer R L, Leach P G L "Adiabatic invariants and Mixmaster catastrophes" *Phys. Rev. D* 57 4691 (1998); gr-qc/ 9712027 †
- Hartle J B, Hawking S W *Phys. Rev. D* 28 2960 (1983); Hawking S W *Nucl. Phys. B* 239 257 (1984)
- Vilenkin A Phys. Lett. B 117 25 (1982); Phys. Rev. D 27 2848 (1983);
 30 509 (1984); 37 888 (1988); Linde A D Zh. Eksp. Teor. Fiz. 87 369 (1984) [Sov. Phys. JETP 60 211 (1984)]; Zel'dovich Ya B, Starobinskiĭ A A Pis'ma Astron. Zh. 10 323 (1984) [Sov. Astron. Lett. 10 135 (1984)]; Rubakov V A Phys. Lett. B 148 280 (1984)
- 19. Landau L D, Private communication
- Al'tshuler B L, Barvinsky A O Usp. Fiz. Nauk 166 459 (1996) [Phys. Usp. 39 429 (1996)]
- 21. Halliwell J J, Hartle J B Phys. Rev. D 41 1815 (1990)
- 22. Barvinsky A O, Kamenshchik A Yu Phys. Rev. D 50 5093 (1994)
- Barvinsky A O, Kamenshchik A Yu, Mishakov I V Nucl. Phys. B 491 387 (1997)
- Landau L D, Lifshitz E M Kvantovaya Mekhanika (Quantum Mechanics) (Moscow: Nauka, 1974) [Translated into English (Oxford: Pergamon Press, 1977)]
- 25. Parker L, Fulling S A Phys. Rev. D 7 2357 (1973)
- Starobinskiĭ A A Pis'ma Astron. Zh. 4 155 (1978) [Sov. Astron. Lett. 4 82 (1978)]
- 27. Barrow J D, Matzner R A Phys. Rev. D 21 336 (1980)
- 28. Hawking S W, in *Relativity, Groups and Topology II* (Amsterdam: North-Holland Physics Publ., 1984)
- 29. Page D N Class. Quantum Grav. 1 417 (1984)
- Kamenshchik A Yu, Khalatnikov I M, Toporensky A V Int. J. Mod. Phys. D 6 673 (1997); gr-qc/9801064
- 31. Kamenshchik A Yu, Khalatnikov I M, Toporensky A V Int. J. Mod. Phys. D 7 129 (1998)
- Khalatnikov I M, Kamenshchik A Yu, Toporensky A V, in Proc. of 5th Course on Current Topics in Astrofundamental Physics, Erice, Italy, September 7–15, 1996 (Eds N Sanchez, A Zichichi) (Singapore: World Scientific, 1997) p. 271
- 33. Khalatnikov I M, Kamenshchik A Yu Phys. Rep. 288 513 (1997)
- 34. Glashow S L, Iliopoulos J, Maiani L Phys. Rev. D 2 1285 (1970)
- 35. Green M B, Schwarz J H, Witten E Superstring Theory Vols 1, 2 (Cambridge, New York: Cambridge Univ. Press, 1987) [Translated into Russian (Moscow: Mir, 1990)]; Brink L, Henneaux M Principles of String Theory (New York: Plenum Press, 1988) [Translated into Russian (Moscow: Mir, 1991)]; Polyakov A M Gauge Fields and Strings (New York: Harwood Academic Publ., 1987) [Translated into Russian (Moscow: Landau Institute of Theoretical Physics, 1995)]
- 36. Polyakov A M Phys. Lett. B 103 211 (1981)
- Linde A D Fizika Élementarnykh Chastits i Inflyatsionnaya Kosmologiya (Particle Physics and Inflationary Cosmology) (Moscow:

[†]See, for example, http://xxx.lanl.gov

Nauka, 1990) [Translated into English (Chur, Switzerland, New York: Harwood Academic Publ., 1990)]

- Barvinsky A O, Kamenshchik A Yu Class. Quantum Grav. 7 L181 (1990)
- 39. Barvinsky A O Phys. Rep. 230 237 (1993);
- 40. Kamenshchik A Yu Phys. Lett. B 316 45 (1993)
- 41. Barvinsky A O, Kamenshchik A Yu Phys. Lett. B 332 270 (1994)
- 42. Barvinsky A O, Kamenshchik A Yu Int. J. Mod. Phys. D 5 825 (1996)
- 43. Esposito G, Kamenshchik A Yu, Miele G Phys. Rev. D 56 1328 (1997)
- 44. Kamenshchik A Yu, Lyakhovich S L Nucl. Phys. B 495 309 (1997)
- 45. Khalatnikov I M, Mezhlumian A Phys. Lett. A 169 308 (1992)
- 46. Khalatnikov I M, Schiller P Phys. Lett. B 302 176 (1993)
- 47. Amendola L, Khalatnikov I M, Litterio M, Occhionero F *Phys. Rev. D* **49** 1881 (1994)
- 48. Kamenshchik A Yu, Khalatnikov I M, Toporensky A V *Phys. Lett. B* **357** 36 (1995)
- Kamenshchik A Yu, Khalatnikov I M, Toporensky A V Int. J. Mod. Phys. D 6 649 (1997); gr-qc/9801064
- 50. Kasner E M Am. J. Math. 43 217 (1921)
- Gross D J, Wilczek F *Phys. Rev. Lett.* **30** 1343 (1973); Politzer D H *Phys. Rev. Lett.* **30** 1346 (1973)
- Burd A, Buric N, Ellis G F R *Gen. Rel. Grav.* 22 349 (1990); Hobill D, Bernstein D, Welge M, Simkins D *Class. Quantum Grav.* 8 1155 (1991)
- 53. Ramani A, Grammaticos B, Bountis T Phys. Rep. 180 159 (1989)
- Contopoulos G, Grammaticos B, Ramani A J. Phys. A 27 5357 (1994); Demare J, Scheen C J. Phys. A 29 59 (1996); Phys. Lett. B 399 207 (1997)
- Belinskii V A, Khalatnikov I M Zh. Eksp. Teor. Fiz. 63 1121 (1972) [Sov. Phys. JETP 36 591 (1973)]
- Barrow J D, Dabrowski M P "Is there chaos in low-energy string cosmology?" hep-th/9711049
- 57. Gasperini M, Veneziano G Astron. Phys. 1 317 (1993)
- Calzetta E, Thibeault M "Perturbative analysis of Bianchi IX using Ashtekar formalism" gr-qc/9710128
- Ashtekar A Lectures on Non-Perturbative Quantum Gravity (Singapore: World Scientific, 1991)
- 60. Gilmore R Catastrophe Theory for Scientists and Engineers (New York: Wiley, 1981)
- Starobinsky A A "Current trends in field theory, quantum gravity and strings" *Lect. Notes Phys.* 246 (Heidelberg: Springer-Verlag, 1986) p. 107; Kolb E W, Turner M *The Early Universe* (Frontiers in Physics, Vol. 69) (Reading, Mass.: Addison-Wesley, 1990)
- Smoot J Aph. J. 396 L1 (1992); Strukov I A et al. Pis'ma Astron. Zh. 18 387 (1992) [Sov. Astron. Lett. 18 153 (1992)]
- 63. Starobinskiĭ A A *Pis'ma Astron. Zh.* **9** 579 (1983) [*Sov. Astron. Lett.* **9** 302 (1983)]
- Dirac P A M Proc. R. Soc. London Ser. A 246 333 (1958); Lectures on Quantum Mechanics (New York: Belfer Graduate School of Science, Yeshiva University, 1964); DeWitt B S Phys. Rev. 160 1113 (1967); Wheeler J A, in Batelles Rencontres (New York: Benjamin, 1968)
- Coleman S Phys. Rev. D 15 2929 (1977); Callan C G, Coleman S Phys. Rev. D 16 1762 (1977)
- 66. Hawking S W, Page D N Nucl. Phys. B 264 185 (1986)
- Grishchuk L P, Rozhansky L V Phys. Lett. B 208 369 (1988); 234 9 (1990)
- Lavrelashvili G V, Rubakov V A, Serebryakov M S, Tinyakov P G Nucl. Phys. B 329 98 (1990); Rubakov V A, Tinyakov P G Nucl. Phys. B 342 430 (1990)
- Salopek D S, Bond J R, Bardeen J M Phys. Rev. D 40 1753 (1989);
 Fakir R, Unruh W G Phys. Rev. D 41 1783 (1990); Fakir R, Habib S, Unruh W G Aph. J. 394 396 (1992); Fakir R, Habib S Mod. Phys. Lett. A 8 2827 (1993)
- Barvinsky A O, Kamenshchik A Yu, Karmazin I P Phys. Rev. D 48 3677 (1993)
- Belinskii V A, Grishchuk L P, Zel'dovich Ya B, Khalatnikov I M Zh. Eksp. Teor. Fiz. 89 346 (1985) [Sov. Phys. JETP 62 195 (1985)]; Belinskii V A, Khalatnikov I M Zh. Eksp. Teor. Fiz. 93 784 (1987) [Sov. Phys. JETP 66 441 (1987)]
- 72. Bousso R, Hawking S W Phys. Rev. D 52 5659 (1995); 54 6312 (1996)
- 73. Carriga J, Vilenkin A Phys. Rev. D 56 2464 (1997)

- 74. Halliwell J J, Hawking S W *Phys. Rev. D* **31** 1777 (1985); Laflamme R *Phys. Lett. B* **198** 156 (1987)
- Hawking S W Commun. Math. Phys. 55 133 (1977); Dowker J S, Critchley R Phys. Rev. D 16 3390 (1977)
- DeWitt B S Dynamical Theory of Groups and Fields (New York: Gordon and Breach, 1965); Barvinsky A O, Vilkovisky G A Phys. Rep. 119 1 (1985)
- Fradkin E S, Vilkovisky G A Phys. Lett. B 55 224 (1975); Batalin I A, Vilkovisky G A Phys. Lett. B 69 309 (1977); Batalin I A, Fradkin E S Phys. Lett. B 128 303 (1983); J. Math. Phys. 25 2426 (1984); Riv. Nuovo Cimento 9 Ser. 3 (10) 1 (1986); Ann. Inst. Henri Poincare 49 145 (1988)
- Hwang S Phys. Rev. D 28 2614 (1983); Ohta N Phys. Rev. D 33 1681 (1986)
- Dirac P A M Lectures on Quantum Mechanics (New York: Belfer Graduate School of Science Yeshiva University, 1964) [Translated into Russian (Moscow: Mir, 1968)]
- Hanson A, Regge T, Teitelboim C Constrained Hamiltonian Systems (Roma: Accademia Nazionale dei Lincei, 1976); Sundermeyer K "Constrained dynamics: with applications to Yang-Mills theory, general relativity, classical spin, dual string model" Lect. Notes Phys. 169 (Berlin, New York: Springer-Verlag, 1982); Henneaux M, Teitelboim C Quantization of Gauge Systems (Princeton, N.J.: Princeton University, 1992)
- 81. Lifshitz E M Zh. Eksp. Teor. Fiz. 16 587 (1946)
- 82. Arnold V Ann. Inst. Fourier XVI 319 (1966)
- Floratos E G, Iliopoulos J Phys. Lett. B 201 237 (1988); Arakelyan T A, Savvidy G K Phys. Lett. B 214 350 (1988); 223 41 (1989); Bars I, Pope C N, Sezgin E Phys. Lett. B 210 85 (1988); DeWitt B, Marquard U, Nicolai H Commun. Math. Phys. 128 39 (1990)
- Wigner E P Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra (Pure and Applied Physics, Vol. 5) (New York: Academic Press, 1959)
- Starobinsky A A Pis'ma Zh. Eksp. Teor. Fiz. 57 606 (1993) [JETP Lett. 55 489 (1993)]; Sokolov I Yu Pis'ma Zh. Eksp. Teor. Fiz. 57 601 (1993) [JETP Lett. 57 617 (1993)]; Stevens D, Scott D, Silk J Phys. Rev. Lett. 71 20 (1993); Lachièze-Rey M, Luminet J-P Phys. Rep. 254 135 (1995)
- Buchbinder I L, Fradkin E S, Lyakhovich S L, Pershin V D Int. J. Mod. Phys. A 6 1211 (1991)
- 87. Burnett G A Phys. Rev. D 51 1621 (1995)
- Mandelbrot B B *The Fractal Geometry of Nature* (San Francisco: W.H. Freeman, 1982); Zel'dovich Ya B, Sokolov D D Usp. Fiz. Nauk 146 493 (1985) [Sov. Phys. Usp. 28 608 (1985)]
- 89. Starobinsky A A, in *Proc. of 15th Int. Conf. on General Relativity and Gravitation, Pune, India, 1997* (to be published)
- Nikomarov E S, Khalatnikov I M Zh. Eksp. Teor. Fiz. 75 1176 (1978) [Sov. Phys. JETP 48 592 (1978)]
- 91. Giddings S B, Strominger A Nucl. Phys. B 306 890 (1988)
- Abramowitz M, Stegun I A (Eds) Handbook of Mathematical Functions (New York: Dover Publ., 1965) [Translated into Russian (Moscow: Nauka, 1979)]
- Landau L D, Abrikosov A A, Khalatnikov I M Dokl. Akad. Nauk SSSR 95 497, 773, 1177 (1954); 96 261 (1954)
- Landau L D Dokl. Akad. Nauk SSSR 44 339 (1944); Landau L D, Lifshitz E M Gidrodinamika (Hydrodynamics) (Moscow: Nauka, 1988) [Translated into English (New York: Pergamon Press, 1970)]