Physics – Uspekhi **39** (5) 429 – 459 (1996)

ON THE 75TH ANNIVERSARY OF A D SAKHAROV'S BIRTH

PACS numbers: 0.4.60.Ds, 98.80.Bp, 98.80.Cq, 98.80.Hw

Quantum cosmology and physics of transitions with a change of the spacetime signature

B L Al'tshuler, A O Barvinsky

Contents

Abstract. This paper examines elements of the general theory of transitions with changing spacetime signature in quantum gravity and cosmology as suggested in a pioneer work of A D - Sakharov. Unlike the conventional formal method for functional integration, this approach uses as the starting point the Dirac – Wheeler – DeWitt operator quantization and its reduction to quantization in Arnowitt – Deser – Misner variables. It has been demonstrated that motivation to consider Euclidean – Lorentzian transitions consists in global ambiguity of physical reduction on the phase space of the theory (a gravitational analog of the problem of Gribov's copies). This ambiguity in particular results in the indefinite sign of the physical inner product of quantum states and leads to the concept of third quantization. An alternative approach is quantization in the York gauge and in special variables of conformal superspace. This quantization is likely to provide a global solution of the ambiguity problem, but the problems of Euclidean – Lorentzian transitions arise in this formalism in the language of complexification of a conformal space of the Wick-rotation type for the time variable in the Klein – Gordon equation. The problem of origin of the early Universe via gravitational tunnelling in Hartle-Hawking and Vilenkin quantum states is considered to illustrate applications of the general theory. The mechanism is described by which loop effects generate the normalizable distribution function for the ensemble of chaotic inflationary universes. In a model with a large non-minimal coupling constant for the scalar inflaton, this mechanism gives rise to a sharp probability peak at sub- Planckian values for the Hubble constant which is in good agreement with the contemporary observational status of the cosmological inflation theory.

B L Al'tshuler, A O Barvinsky P N Lebedev Physics Institute, Russian Academy of Sciences, Leninskiĭ prosp. 53, 117924 Moscow, Russia
Tel. (7-095) 135 83 39; (7-095) 132 60 49
E-mail: altshul@lpi.ac.ru, barvin@rq2.fian.msk.su

Received 19 February 1996
Uspekhi Fizicheskikh Nauk **166** (5) 459–492 (1996)
Translated by Morozov; edited by S D Danilov

1. Introduction

The theory of gravity and cosmology occupied a special place among versatile scientific interests of A D Sakharov. These issues are covered in 17 papers included in the Complete Works of A D Sakharov published recently (excepting secret reports on the construction of atomic weapons) [1]. They concern induced gravitation, explanation of barionic asymmetry of the Universe, multisheet Universe, cosmological models with the reversal of time arrow, and finally quantum cosmology treated in a paper on 'Cosmological transitions with changing metric signature' which was written in exile in 1984 [2]. (Some fundamental works of A D Sakharov were also published in a special issue of *Uspekhi Fizicheskikh Nauk* (*Soviet Physics – Uspekhi*) on the occasion of his seventieth birthday [3]).

The present paper is not designed to be a comprehensive review although the Introduction lists a few selected topical problems and approaches of gravity quantization. Science has made considerable progress during 12 years that elapsed after A D Sakharov had published his paper [2]. ("There are miraculous achievements in science. Although I do not believe in the possibility to have a universal theory in the near future (or to have it at all), I do see staggering, fantastic progress even within my lifespan. I hope this onward march of science will never stop but will be continuously spreading and growing.'' — A D Sakharov in [4], p 22). In fact, we think that the 'staggering, fantastic progress' in quantum gravity is difficult to expect in the near future. Nevertheless, we try to discuss some ideas of A D Sakharov's paper [2] with regard to the current state of this discipline and to give more or less rigorous mathematical formulation for selected heuristic hypotheses of the author.

How to combine the general theory of relativity (GTR) and quantum physics has been eagerly sought by a few generations of physicists. But despite all rapid developments that quantum gravity and quantum cosmology, its natural application, have experienced during the last 10 – 20 years, a number of most important conceptual problems remain to be solved. This is not without reason. Such terms as 'time', 'the past and the future', 'causality', 'unitarity', 'reference frames', 'event' are used to designate clear-cut physical notions or basic principles both in GTR and in the quantum theory although each of these words needs to be additionally defined as soon as one begins to quantize the metric tensor, that is when spacetime intervals are considered to be quantum objects.

Assuming quantum fluctuations of the metric tensor and following strong analogies with the string theory, one can make a further step, that is include manifolds of different topology, dimensions and signatures in the quantum ensemble. Quantum functional averaging over spacetime diffeomorphisms ('reference frame quantization') 'spreads' gauge-noninvariant objects which are represented in the framework of GTR by any observables with a limited spacetime support. The standard procedure of gauge fixation eliminates 'spreading' but makes quantum predictions gauge-dependent. At the same time, the use of only gauge-invariant objects (topological numbers, integrals over the entire spacetime...) raises the problem of the description of local geometry in this language.

The requirement of invariance of quantum observables and quantum 'wave function of the Universe' under general coordinate transformations means imposition of constraints. In particular, time altogether drops out of the quantum description of gravitation since all dynamical evolution can be mimicked by time reparametrization. This well-known problem of time in quantum gravity raises very important questions: How can physical time be introduced? Is it possible to do this unambiguously, taking a certain subsystem of the quantum system as the 'clock'? What does it mean to quantize a system in which all observations are from 'the inside' and only relative probabilities make sense? These and related questions have lately given rise to special fields of research. We shall briefly discuss some quantization methods and trends of reasoning, referring the reader to original publications and reviews whenever possible.

1.1 Canonical quantization

Canonical, non-covariant quantization of Einstein's GTR has been developed in pioneering works of Dirac, Wheeler, and DeWitt [5, 6, 7] (see, for instance, review [8]). The most important result of quantization is a formal loss of spacetime; the wave function satisfying the Wheeler – DeWitt constraint equation (00-component of the Einstein equations)

$$\hat{H}_{\mathbf{x}} \Psi(q^i) = 0, \tag{1.1}$$

depends only on the superspace point:

$$q^i \equiv \{g_{ab}(\mathbf{x}), \varphi(\mathbf{x})\} \tag{1.2}$$

and is time-independent. Both the three-dimensional metric $g_{ab}(\mathbf{x})$ and the fields of matter $\varphi(\mathbf{x})$ on this three-dimensional manifold constitute the primary language of the theory. It is *generally accepted* that time (hence, 4-dimensionality) and the light cone appear in this approach only in the semiclassical approximation for the solution of Eqn (1.1).

However, it is possible to choose, from the very beginning, a more natural physical way and postulate that only physical degrees of freedom are subject to quantization (for free gravitation in four dimensions, there are only two graviton polarizations; in three dimensions, the number of physical local gravitational degrees of freedom is equal to zero). This unitary gauge approach based on the Arnowitt – Deser – Misner (ADM) reduction to physical variables [9] has been thoroughly examined in Ref. [8]. A large part of the present paper is also concerned with further development of this approach. Despite the fact that in the ADM quantum reduction method, time, quantum-mechanical scalar product, conserved probabilities, and physical Hilbert space are introduced without any reference to the semiclassical regime, there are intrinsic difficulties and problems: How much do physical predictions depend on the choice of the gauge condition (the choice of the reference frames)? What is the physical meaning of emerging Gribov's copies? These problems are discussed below (Section 3).

In the framework of the operator quantization method examined in the present paper, with superspace as the language, the possibility to use one of the variables q^i (1.2) or their combination as the classical clock fixed by the gauge condition is traditionally associated with the fact that the signature of one of the superspace axes is negative (see Section 2). However, any 'monotonically changing' coordinate can play the part of the clock. (Words in inverted commas have sense only after the shape of the Hamiltonian is given; by definition, 'monotony' is broken when the Poisson bracket of a given quantity with the Hamiltonian vanishes). The signature of superspace is *a priori* unrelated to the spacetime signature. The problem is that the gravitational analog of Gribov's copies inevitably exists at superspace caustics which preclude the global use of the gauge that fixes monotonically changing time over the entire superspace q^i. This is actually the main incentive for the transition to third quantization. In fact, a large part of the present paper is devoted to this and related problems.

1.2 Covariant BVF quantization

A link between the canonical and covariant methods of gravity quantization is constituted by the Batalin – Vilkovisky – Fradkin (BVF) quantization [10] the relation of which to the Dirac – Wheeler – DeWitt quantization and unitary reduction has been discussed in detail in Ref. [8]. The BVF quantization uses relativistic gauge when the Lagrange multipliers become quantum dynamical variables; concurrently, the Faddeev – Popov ghosts need to be introduced to compensate extra degrees of freedom, and the condition of the vanishing Becchi – Rouet – Stora – Tyutin (BRST) charge is imposed upon vectors of the extended Hilbert space (see Ref. [11]). This approach is very convenient for the construction of S-matrix which, as can be proved in general form, is independent of the gauge on the mass shell. However, in the cosmological context, where there is no such notion as the in-limit at the asymptotically flat infinity of spacetime, this method is either reduced to the Dirac – Wheeler – DeWitt quantization or implies other types of operator quantization which remain to be constructed.

1.3 Integration over trajectories

"However the split into three spatial dimensions and one time dimension seems to be contrary to the whole spirit of relativity. Moreover, it restricts the topology of spacetime to be the product of the real line with some three-dimensional manifold, whereas one would expect that quantum gravity would allow all possible topologies of spacetime including those which are not products". This is a statement of Hawking from his criticism of canonical quantization in a paper for the book [12] issued on the occasion of Einstein's centenary. In the covariant (non-canonical) approach, the principal tool is the Feynman functional integral over spacetime fields ('trajectories'):

$$K(q, q') = \int \exp\left\{ - \int_{q'(\partial M')}^{q(\partial M)} \sqrt{g}\, L\, \mathrm{d}^4 x \right\} [\mathrm{D} g_{\mu\nu} \mathrm{D}\varphi] . \qquad (1.3)$$

Here, L is Lagrangian, $g = \det g_{\mu\nu}(x, t)$, the gauge and the measure are included in symbol $[\mathrm{D}g\, \mathrm{D}\varphi]$; q', q are the fields (1.2) on the hypersurfaces $\partial M', \partial M$ bounding the 4-manifold M. The covariant formula (1.3) is in agreement with the canonical quantization method since the kernel $K(q, q')$ is the solution of the Wheeler – DeWitt equation (1.1) with respect to each variable q, q' if the functional integral in (1.3) is defined properly [13 – 16]. Specifically, Hartle and Hawking obtained their celebrated solution of the Wheeler – DeWitt equation, the so-called 'no-boundary' wave function of the Universe, by postulating the Euclidean form, shrinking the initial hypersurface $\partial M'$ to a point, and requiring all the fields at this point to be regular [17, 18]. For all that, the analogy between the quantum-gravitational kernel (1.3) and the Feynman kernel of quantum mechanics or quantum field theory should be used with great caution. The standard composition law used in Ref. [17] and based on the integration over all end points at a given moment is inapplicable in this case because $K(q, q')$ satisfies constraint equations, and the naive integral over q is infinite (see discussion in Section 2.3).

1.4 The effective action method

Instead of the amplitude (1.3), it is possible to formulate the theory in the language of the explicitly covariant effective action $\Gamma(\bar{g}_{\mu\nu}, \bar{\varphi})$ defined by the Legendre transformation of the generating functional of external currents. In the tree approximation, Γ coincides with the initial classical action, then the one-loop term of the $\hbar$ order follows, etc. An advantage of the effective action method is that it does not deal with quantum states but yields covariant algorithm, i.e. allows covariant regularization for quantum averages (to be shown in Section 4, where this formalism is applied to the problem of quantum birth of the Universe and makes it possible to calculate the distribution function and to obtain the inflationary scale). At the same time, the effective action method cannot replace basic principles of operator quantization. The form of Γ depends on the choice of field asymptotics in the functional integral (i.e. on the choice of the initial and final states). Moreover, in the case of the standard definition of the effective action, results of the calculation depend upon parametrization of fields — arguments of the functional integral. Off mass shell, Γ depends on the choice of gauge-fixing condition. An important step to constructing the unambiguous effective action has been made by Vilkovisky and DeWitt based on an excellent idea of geometrization of the space of 'trajectories' [19, 20]. (See also Refs [21 – 24] and references therein for applications of this method.)

The functional representation of the Vilkovisky – DeWitt effective action in Refs [23, 24] accounts for the so-called 'residual', 'large' diffeomorphisms (Gribov's copies — 'zeroes' of the Faddeev – Popov operators of the Landau – DeWitt relativistic gauge) which is physically equivalent to the requirement of invariance of observables with respect to these diffeomorphisms. It has been shown that this requirement is non-trivial, imposes restrictions on the admissible background metric $\bar{g}_{\mu\nu}$, and leads to the 'quantum repulsion' from 'free' (without external sources) solutions of classical dynamical equations. The environment and the external sources arise in different contexts when one attempts to obtain the rational classical limit in quantum cosmology.

1.5 Observational status of quantum gravity. The problem of classical limit

One may think of two objections to the apparently explicit and well-known statement that quantum gravity cannot be (and will not be in the near future) experimentally verified since the Planck scale $\sim 10^{-33}$ is absolutely unattainable.

The first 'experimental site' is the early Universe, the problem of the initial state. Similar to inflation theories that establish close relations between theoretical ideas about superhigh energy scales of Grand Unification and microscopical structural features of the observed Universe, the problem of choice of the quantum initial state of the Universe may likewise prove to be of practical importance.

Another 'test' for quantum cosmology is our everyday experience, the indisputable fact of our living in the practically classical $(3+1)$-dimensional Riemann space in which spacetime distances are sufficiently well-determined due to the smallness of quantum metric dispersion. If the theory encounters difficulty in 'preparing' such a narrow-peaked coherent quantum state of the Universe, it (the theory) must be replaced and the basic principles must be revised, if necessary. The problem of classical limit in quantum gravity has become a serious test for the theory in question and gave a new impetus to the re-examination of fundamentals of the quantum theory (see Refs [25, 26] and references therein). Almost all the authors concerned with the problem of classical limit argue that neither time nor classical spacetime can be constructed from the pure wave function of the Universe since the system must be open and 'external' degrees of freedom are needed averaging over which would nullify non-diagonal elements of the density matrix. However, it can not be excluded that the general concern about basic principles is unwarranted. It is universally known that in quantum mechanics, the semiclassical nature of the wave function is not sufficient for a particle to exhibit 'classical' behaviour. It requires the packet, the coherent state, the pre-exponential multiplier whose maximum follows the classical trajectory given by the Hamilton-Jacobi equations while the dispersion about this trajectory must be small as well as the spreading rate of the packet. Therefore, the problem of the appearance of the approximately classical spacetime is the problem of preparing the initial state of the Universe with small dispersion of metric tensor components near the solution of the Hamilton-Jacobi equation. Section 4 demonstrates that in the quantum creation of the Universe (tunnelling from the state with the Euclidean spacetime signature), such an initial state may occur if loop quantum gravitational corrections are taken into account. Parameters of this initial state agree with the observational status of the inflationary Universe.

Also, it is worth noting that the popular requirement of decoherence of various semiclassical Everett's branches of the Universe evolution may prove superfluous because we (observers) evidently belong to the same branch and actually see everything from the inside; that is, we perform the necessary reduction of the wave packet by our own existence.

1.6 A change of topology

Three-dimensional hypersurfaces ∂M, $\partial M'$ (boundaries of the manifold M in (1.3)) are not necessarily simply-connected. It is quite possible to consider the 'multiparticle' ('multi-Universe') state

$$q(\partial M) = \left\{ g_{ab}^{(1)}(\mathbf{x}_1)\,\varphi^{(1)}(\mathbf{x}_1) \ldots g_{ab}^{(n)}(\mathbf{x}_n), \varphi^{(n)}(\mathbf{x}_n) \right\}, \qquad (1.4)$$

where $q(\partial M)$ is a set of field values on various 3-manifolds the number of which in $\partial M'$ may generally speaking be different. The idea of possible topological transitions ('spacetime foam') in quantum gravity was first suggested by Wheeler [27]. The inclusion of manifolds with complicated topology in the arguments of kernel (1.3) and the description of the Mega-Universe naturally requires transition to third quantization when the wave function of the Universe $\Psi(q)$ satisfying the Wheeler – DeWitt equation (1.1), i.e. the infinite number of the Klein – Gordon-type equations (equal the number of points in the 3-space), is taken as the operator. New physics arises in processes with the changing number of universes in (1.3) which evidently requires the introduction of interactions between universes ('relativistic particles'), i.e. the insertion of nonlinear in $\Psi(q)$ terms into Eqn (1.1) [28, 29]. The key questions are: What is responsible for this interaction? Is it possible to extract it from the Einstein action, the initial action of quantum gravity? The most natural way (also ensuing from analogy to the string theory) seems to determine the vertices of this interaction of the order N as follows: to cut out n 4-'circles' from the Euclidean 4-sphere and then to perform functional averaging over all 4-geometries topologically equivalent to a sphere, with fixed values of 3-geometries at the boundary of each cut 'circle' [17]. These distant 4-regions may be considered as wormhole throats joining the Euclidean 4-sphere and other Euclidean spaces, or Lorentzian universes, in the case of changing the signature. (The no-boundary Hartle-Hawking wave function is the 'vertex' of order 1; the creation of the Lorentz Universe 'from nothing' corresponds to a change of the signature at the maximum radius of the 'cutting' which is in this case a hemisphere (see Fig. 3 and the discussion in Section 2)). However, the amplitudes thus determined in Ref. [17] must satisfy the Wheeler-DeWitt equation with respect to each of their 'tails' due to functional averaging over lapse and shift functions; that is, in the language of the S-matrix, they are localised on the 'mass shell' and can not therefore serve as 'bare' vertices.

A special line of research is the theory of wormholes and the development of a challenging idea of renormalization [30] and, possibly, dynamical fixation of all constants of low-energy physics due to the presence, at each point of a large 'parent' universe, of virtual wormholes which connect this point with another point in the same universe or with a 'daughter' universes. In order to illustrate this situation, let us give expressions for the metric of the simplest wormhole in n-dimensional Euclidean space in three different parametrizations of the radial coordinate:

$$\mathrm{d}s^2 = \mathrm{d}r^2 + (r^2 + b^2)\,\mathrm{d}\Omega^{(n-1)} \qquad (1.5a)$$

$$= \left(1 + \frac{b^2}{4z^2}\right)^2 \left[\mathrm{d}z^2 + z^2\,\mathrm{d}\Omega^{(n-1)}\right] \qquad (1.5b)$$

$$= \frac{R^2}{R^2 - b^2}\,\mathrm{d}R^2 + R^2\,\mathrm{d}\Omega^{(n-1)}. \qquad (1.5c)$$

Here, $\mathrm{d}\Omega^{(n-1)}$ is an element of $(n-1)$-dimensional sphere with unit radius;

$$r = z - \frac{b^2}{4z} = \sqrt{R^2 - b^2}\,. \qquad (1.6)$$

The expression (1.5a) is especially illuminating: at $b = 0$, there is a flat metric in the spherical coordinate system whereas at any arbitrarily small $b \neq 0$ there are two asymptotically flat infinite spaces joined by a throat of radius b. Version (1.5b) has been used in Ref. [31] which postulates the loss of coherence by an initially 'pure' quantum-mechanical state due to irreversible leakage of information through wormholes. Quite different hypothesis concerning coherent effects of wormholes has been suggested in Ref. [30]. In a recent paper by Rubakov [32], the role of quantum transitions with changing topology has been studied in the framework of the string theory, i.e. in two-dimensional gravitation. The author considers creation of 'daughter' universes (strings) to be nothing but standard emission of gravitons in D-dimensional spacetime ('target space'). Also, it has been shown in Ref. [32] that both phenomena (the loss of coherence and coherent renormalization of phenomenological action constants, i.e. fields in 'the target space') take place, each in its own energy region.

It is essential that the action

$$S \sim \int R\sqrt{g}\,\mathrm{d}^n x \sim b^{n-2}\,,$$

calculated for metric (1.5) is small for small b, i.e. quantum dynamics does not significantly suppress the appearance of microscopic wormholes in each point of the 'large' Universe. The theory of renormalization of constants by wormholes may formally be reduced to the standard diagram technique in which specific supernonlocal (independent of 'inlet' and 'outlet' points of a wormhole) Green's functions correspond to lines while the vertices are represented by integrals of various local operators over the entire volume. In the dilute gas (non-interacting wormholes) approximation, this effectively leads to the renormalization of observed constants. Here, the main contribution comes from the deep infrared region rather than the ultraviolet one (unlike the situation in the standard quantum field theory), and the world's properties at large distances really influence the observed local physics.

In other language, the dilute gas approximation corresponds to the application of the mean field approximation to the Wheeler – DeWitt equation with nonlinear terms for the third quantized wave function of the Universe. It should be emphasized that these words and notions hardly understandable in the context of 4-dimensionality become sufficiently clear and unambiguous as soon as one discusses the string theory (two-dimensionality) [33 – 35].

Today, extensive studies on classical and quantum wormholes are underway (see Refs [36, 37] containing many useful references). Speaking of wormholes, one cannot help mentioning another recent line of research the key words of which are 'Lorentzian wormholes' and 'time machine', with all accompanying questions like 'What happens if someone kills his grandmother?' inseparable from the backward journey in time [38, 39].

It is worth noting close relationship between wormholes and the signature change in the radial coordinate which is most obvious if the wormhole metric is written in the form (1.5c), where R is the scale factor of the $(n-1)$-sphere. Clearly, the coordinate R becomes timelike at $R < b$ (the classically forbidden region for metric (1.5)). It is this pole expression for the g_{00}-component of the metric that was used by A D Sakharov to illustrate the idea of changing signature (see below, (1.7)).

1.7 A change of signature

In Ref. [2], A D Sakharov suggested that Euclidean and pseudo-Euclidean geometries with the different number of time-axes should be considered on equal basis in the functional integral (1.3); in this case, the imaginary unity before the action emerges automatically due to the multiplier $\sqrt{g}$ in the exponent of (1.3) for the Lorentzian signature (the odd number of time axes, $g < 0$). The idea of additional time-dimensions had been suggested earlier (see references in Ref. [75] and comments by I D Novikov, V P Frolov, V A Rubakov, I Ya Aref'eva, and I V Volovich on [2] in [1], pp 309 – 313).

A D Sakharov illustrated the central idea of Ref. [2], the possibility of quantum tunnelling transitions between spaces with different signature, by a simple example of the sign alteration of the g_{00}-component of the metric tensor upon crossing the pole at the coordinate x_0, the boundary between the Euclidean and Lorentzian regions:

$$g_{00} = \frac{l}{x_0 - c} \,. \tag{1.7}$$

At $x > c$ and $x < c$ one has the Euclidean and Lorentzian metrics, respectively. According to Sakharov, the proper time determined by the transformation of coordinates:

$$\begin{aligned} &\text{at} \quad x_0 > c \quad x_0 - c = \frac{\tau^2}{4l}, \qquad g_{00} \to g'_{00} = 1 \,, \\ &\text{at} \quad x_0 < c \quad c - x_0 = \frac{t^2}{4l}, \qquad g_{00} \to g'_{00} = -1 \,. \end{aligned} \tag{1.8}$$

can be introduced in either world. It will be shown in Section 3 that the pole form of the metric (1.7) naturally arises at the bounce point, i.e. at the boundary of the Euclidean-Lorentzian transition in the system of coordinates corresponding to the parametrization of time by the scale factor of the Universe. The appearance of singularity of the (1.7) type has been illustrated by formula (1.5) above. Unlike the wormhole (1.5), the de Sitter Universe with the bounce is the Lorentzian Universe at large values of the scale factor and the Euclidean Universe at small ones.

The present paper discusses at length only the quantum-cosmological aspect of changing spacetime signature, namely with regard to the problem of the creation of the de Sitter Lorentzian Universe from the Euclidean 4-sphere. The consideration of the components of metric tensor as complexified dynamical variables is presented in Refs [41, 42]. Ref. [42] attempts on formulating a dynamical principle for the prediction of both the dimensionality and the signature $(-+++)$ of our Universe. Also, many studies concern solutions for classical equations affecting signature (see Refs [43 – 45] and references therein). The authors of Ref. [46] treated spacetime manifolds with different signatures in the supermembrane language.

1.8 Fundamentally different approaches to quantization of gravity

In the discussion of gravity quantization methods above, we have assumed the Einstein action (or other gravitational actions) to be the primary initial element of the theory. However, a quite different approach (going back to Sakharov's idea of induced gravitation [47] and maximally realized in the string theory [48, 49]) turns out to be most promising. According to this standpoint, both the gravitational action and the spacetime continuum itself are likely to arise only in the phenomenological, long wavelength limit of a more fundamental theory. This revolutionary concept of spacetime and gravitation unlimited by the expansion in the perturbation theory has much in common with Regge's calculus, the dynamical triangulation method, matrix models, the use of the Ashtekar variables, and the treatment of quantum gravity as the topological quantum field theory (see, for instance, the special issue of *Journ. Math. Phys.* [50]). So far, the main objective of these studies has been to define such fundamental notions and quantities as topological vicinity, Hausdorff dimensionality, introduction of the concept of distance in the scaling limit, etc. None of these approaches has yet been applied to cosmology.

Detailed bibliography on quantum gravity up to 1990 is included in Ref. [51] (see also a recent Isham's review lecture [52]). Quantum gravity has been discussed at length at the 1st International Sakharov Conference on Physics (P N Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 21 – 31 May, 1991, see Ref. [53]). We hope that this field of research will be equally well represented at the 2nd Sakharov Conference to be held in Moscow in May, 1996.

In the present paper we confine ourselves to examining the most traditional canonical ADM quantization of the Einstein theory. In Section 2, we introduce constraint equations, investigate peculiarities of the inner product of wave functions satisfying these equations, and show the inevitability of the problem of time and transition to third quantization in quantum gravity within the Dirac – Wheeler – DeWitt method. In Section 3, quantization is performed by the ADM reduction, i.e. in the unitary gauge when four arbitrary functions of the superspace coordinates parametrized by number t are excluded from the quantum ensemble. Section 3 also deals with the problem of Gribov's copies and the possibility to use gauges in the absence of this problem which may, in principle, eliminate the necessity of transition to third quantization. Section 4 is devoted to the construction, in the one-loop approximation, of the wave function of our Universe 'created' from the Euclidean region, i.e. from the state with positive-definite signature. In the Conclusions, we discuss possible lines of further research as well as the relationship between 'anthropic' and 'dynamical' approaches to the explanation of the properties of the observed Universe.

2. Gravitational constraints, the problem of time, and third quantization

We begin this Section with a brief discussion of the special features of Einstein's relativity which give rise to the well-known problem of time and actually provide the basis for the theory of quantum transitions with the changing spacetime signature. The canonical action of the theory on the phase space of three-dimensional metric coefficients g_{ab}, matter fields φ, and momenta conjugate to them p^{ab}, p_{φ} $(a, b, \ldots = 1, 2, 3)$ has the form [6, 9, 54]

$$\boldsymbol{S}[g_{ab},\, p^{ab},\, \varphi,\, p_{\varphi},\, N^{\perp},\, N^{a}] = \int \mathrm{d}t\; \mathrm{d}^3x\, [p^{ab}\dot{g}_{ab} + p_{\varphi}\dot{\varphi} - N^{\perp}H_{\perp} - N^{a}H_{a}]. \tag{2.1}$$

Lapse ($N^{\perp} = (-{}^4g^{00})^{-1/2}$) and shift ($N^a = g^{ab\,4}g_{0b}$) functions enter this expression without time derivatives and may therefore be regarded as Lagrange multipliers in gravitational constraints: the superhamiltonian and supermomenta

$$H_{\perp}(g_{ab},\, \varphi,\, p^{ab},\, p_{\varphi}) = \frac{1}{2}\, G_{ab,\,cd}\, p^{ab}p^{cd} - \sqrt{g}\; {}^3R + H_{\perp}^{\mathrm{mat}}(g_{ab},\, \varphi,\, p_{\varphi}), \tag{2.2}$$

$$H_{a}(g_{ab},\, \varphi,\, p^{ab},\, p_{\varphi}) = -2g_{ab}\nabla_{c}p^{bc} + H_{a}^{\mathrm{mat}}(g_{ab},\, \varphi,\, p_{\varphi}), \tag{2.3}$$

where $G_{ab,\,cd} = g^{-1/2}(g_{ac}g_{bd} + g_{ad}g_{bc} - g_{ab}g_{cd})$ is the local matrix of the DeWitt supermetric, while $H_{\perp}^{\mathrm{mat}}$ and H_{a}^{mat} are the contributions of the matter fields whose specific form depends on the choice of matter. This action gives rise to canonical equations of motion for the coordinates and momenta, by varying it with respect to phase-space variables, while the variation of the Lagrange multipliers imposes non-dynamical constraint equations

$$H_{\perp} = 0, \qquad H_a = 0\,, \tag{2.4}$$

and leaves $N^{\perp}$ and N^a undefined. This arbitrariness is equivalent to the invariance of the gravitational action with respect to 4-dimensional diffeomorphisms. For phase variables, they are generated by canonical transformations with gravitational constraints $(H_{\perp}, H_a)$ as the generators of local infinitesimal diffeomorphisms, normal ($\perp$) and tangent (a) to the constant time surfaces, respectively. Constraints (2.2), (2.3) belong to the first class, according to Dirac's terminology [5], and are in involution with respect to the Poisson bracket $\{\ldots, \ldots\}$ on the phase space [6, 9, 54]

$$\begin{aligned}
\big\{H_{\perp}(\mathbf{x}),\, H_{\perp}(\mathbf{x}')\big\} &= g^{ab}(\mathbf{x})\, H_b(\mathbf{x})\, \partial_a\delta(\mathbf{x},\, \mathbf{x}') - (\mathbf{x} \leftrightarrow \mathbf{x}')\,,\\
\big\{H_{\perp}(\mathbf{x}),\, H_a(\mathbf{x}')\big\} &= -H_{\perp}(\mathbf{x}')\, \partial_a\delta(\mathbf{x}',\, \mathbf{x})\,, \qquad (2.5)\\
\big\{H_a(\mathbf{x}),\, H_b(\mathbf{x}')\big\} &= H_b(\mathbf{x})\, \partial_a\delta(\mathbf{x},\, \mathbf{x}') - (a,\, \mathbf{x} \leftrightarrow b,\, \mathbf{x}')\,.
\end{aligned}$$

The Hamiltonian of a closed cosmological model with the action (2.1) is a linear set of constraints and therefore vanishes on their solutions. This feature is responsible for the well-known problem of time in both classical and quantum cosmologies. At the classical level this problem shows up in the specific properties of the Hamilton – Jacobi gravitational function $\boldsymbol{S}(t,\, g_{ab},\, \varphi;\, t',\, g'_{ab},\, \varphi')$, i.e. action (2.1), calculated on a classical extremal joining configurations $(g'_{ab},\, \varphi')$ and $(g_{ab},\, \varphi)$ taken at the initial (t) and final (t') moments. Due to gravitational constraints, function $\boldsymbol{S}(t,\, g_{ab},\, \varphi;\, t',\, g'_{ab},\, \varphi')$ satisfies the system of Einstein-Hamilton-Jacobi equations derived form (2.4) by substituting functional gradients of S for the momenta

$$H_{\perp}\left(g_{ab},\, \varphi,\, \frac{\delta\boldsymbol{S}}{\delta g_{ab}},\, \frac{\delta\boldsymbol{S}}{\delta\varphi}\right) \equiv \frac{1}{2}\, G_{ab,\,cd}\, \frac{\delta\boldsymbol{S}}{\delta g_{ab}}\, \frac{\delta\boldsymbol{S}}{\delta g_{cd}} - \sqrt{g}\, {}^3R + H_{\perp}^{\mathrm{mat}}\left(g_{ab},\, \varphi,\, \frac{\delta\boldsymbol{S}}{\delta\varphi}\right) = 0, \tag{2.6}$$

$$H_{a}\left(g_{ab},\, \varphi,\, \frac{\delta\boldsymbol{S}}{\delta g_{ab}},\, \frac{\delta\boldsymbol{S}}{\delta\varphi}\right) \equiv -2g_{ab}\nabla_c\, \frac{\delta\boldsymbol{S}}{\delta g_{bc}} + H_{a}^{\mathrm{mat}}\left(g_{ab},\, \varphi,\, \frac{\delta\boldsymbol{S}}{\delta\varphi}\right) = 0, \tag{2.7}$$

Hence, the ordinary Hamilton – Jacobi equation for $\boldsymbol{S}(t,\, g_{ab},\, \varphi;\, t',\, g'_{ab},\, \varphi')$ is reduced to the independence of this function of t (and t')

$$\boldsymbol{S}(t,\, g_{ab},\, \varphi;\, t',\, g'_{ab},\, \varphi') = \boldsymbol{S}(g_{ab},\, \varphi;\, g'_{ab},\, \varphi')\,. \tag{2.8}$$

Thus, the Hamilton – Jacobi function does not bear explicit information about the time characterizing gravitational system dynamics.

At the quantum level, the problem of time arises in the Shrödinger picture of the Dirac – Wheeler – DeWitt quantization. This picture is essentially as follows. In contrast to what is the case of the conventional non-gauge theory, gravitational phase variables in the theory of gravity are not independent and satisfy the constraints (2.4). *A priori*, in the quantum theory, there is a freedom to take this fact into consideration in two different ways. On the one hand, the constraints may be regarded as equations on operators g_{ab}, p^{ab}, φ, p_{φ} which become (as they do in the classical theory) dependent and satisfying special (Dirac's) commutation relations compatible with the constraints [9, 55, 10]. On the other hand, it is possible to consider the operators of phase-space variables as independent and to impose the constraints as equations selecting physical states $|\boldsymbol{\Psi}\rangle$ in full representation space of these operators [5, 6]

$$\hat{H}_{\perp}|\boldsymbol{\Psi}\rangle = 0, \qquad \hat{H}_a\,|\boldsymbol{\Psi}\rangle = 0. \tag{2.9}$$

The second approach is in fact Dirac's quantization scheme in which *independent* canonical operators satisfy the Heisenberg commutation relations. In the coordinate representation $(|\boldsymbol{\Psi}\rangle = \boldsymbol{\Psi}\,[g_{ab}(\mathbf{x}),\, \varphi(\mathbf{x})])$ of commutation relations

$$\begin{aligned}
&\hat{g}_{ab}(\mathbf{x}) = g_{ab}(\mathbf{x}), \qquad \hat{\varphi}(\mathbf{x}) = \varphi(\mathbf{x}),\\
&\hat{p}^{ab}(\mathbf{x}) = \frac{\hbar}{\mathrm{i}}\, \frac{\delta}{\delta g_{ab}(\mathbf{x})}, \qquad \hat{p}_{\varphi}(\mathbf{x}) = \frac{\hbar}{\mathrm{i}}\, \frac{\delta}{\delta\varphi(\mathbf{x})}
\end{aligned} \tag{2.10}$$

Eqns (2.9) assume the form

$$''\left\{-\frac{\hbar^2}{2}\, G_{ab,\,cd}\, \frac{\delta^2}{\delta g_{ab}\delta g_{cd}} - \sqrt{g}\; {}^3R + H_{\perp}^{\mathrm{mat}}\left(g_{ab},\, \varphi,\, \frac{\hbar}{\mathrm{i}}\, \frac{\delta}{\delta\varphi}\right)\right\}'' \times \boldsymbol{\Psi}\,[g_{ab},\, \varphi] = 0, \tag{2.11}$$

$$''\left\{-2\, \frac{\hbar}{\mathrm{i}}\, g_{ab}\nabla_c\, \frac{\delta}{\delta g_{bc}} + H_{a}^{\mathrm{mat}}\left(g_{ab},\, \varphi,\, \frac{\hbar}{\mathrm{i}}\, \frac{\delta}{\delta\varphi}\right)\right\}''\, \boldsymbol{\Psi}\,[g_{ab},\, \varphi] = 0, \tag{2.12}$$

where the double primes indicate that given functional differential operators imply symbolic operator realization of the classical constraints (2.2), (2.3) which implies both the ordering of non-commuting factors and possible quantum corrections proportional to $\hbar$.

The relations (2.11), (2.12) are basic dynamical equations for the physical states in canonical quantum gravity and cosmology, the former being known as the Wheeler – DeWitt equation. Their direct corollary is the time-independence of physical states

$$i\hbar \frac{\partial}{\partial t} |\boldsymbol{\Psi}\rangle = \int d^3x (N^{\perp}\hat{H}_{\perp} + N^a \hat{H}_a) |\boldsymbol{\Psi}\rangle = 0, \qquad (2.13)$$

which is a direct analog of the property (2.8) of the classical Hamilton – Jacobi function.

At first sight, this should lead to a paradoxical conclusion that the system is totally lacking dynamics because the naively calculated quantum averages of any Shrödinger operator $\hat{O}$, $\langle \boldsymbol{\Psi} | \hat{O} | \boldsymbol{\Psi} \rangle$ do not display time-dependence. Actually, the naivety of such calculations follows from the fact that the quantum average itself and the corresponding scalar product of physical states

$$\langle \boldsymbol{\Psi}_1 | \boldsymbol{\Psi}_2 \rangle = \int \prod_{\mathbf{x}} dg_{ab}(\mathbf{x})\, d\varphi(\mathbf{x})\, \boldsymbol{\Psi}^*_1[g_{ab}, \varphi]\, \boldsymbol{\Psi}_2[g_{ab}, \varphi] \qquad (2.14)$$

are poorly defined (diverge) owing to quantum constraints (2.11), (2.12). In fact, these linear homogeneous differential equations on (g_{ab}, φ) imply that their solutions can not be square-integrable (for example, Eqn (2.12) means that the wave function is constant along the orbits of three-dimensional spatial diffeomorphisms in superspace of variables (g_{ab}, φ) and integration over the corresponding directions in this infinite space is likely to diverge). All this indicates that the formalism of the aforementioned Dirac – Wheeler – DeWitt quantization is not closed: it does not explicitly contain time and has no physically meaningful inner product of quantum states. It is natural to suggest that these difficulties result from the fact that physical time is hidden (parametrized) among variables of the superspace of 3-metrics and matter fields or, in a broader sense, among variables of the entire phase space. This hypothesis is supported by additional arguments concerning the structure of the Wheeler-DeWitt equation (2.11). The kinetic term in the Hamiltonian constraint,

$$\frac{1}{2} G_{ab,cd} p^{ab} p^{cd} = \frac{1}{\sqrt{g}} \left(p_{ab} p^{ab} - \frac{1}{2} p^2 \right),$$
$$p \equiv g_{ab} p^{ab}, \qquad p_{ab} \equiv g_{ac} g_{bd} p^{cd}, \qquad (2.15)$$

is indefinite, and the DeWitt supermetric has the hyperbolic signature

$$\operatorname{Sign} G_{ab,cd} = (-+++++) \qquad (2.16)$$

with the 'timelike' sign in the conformal mode sector of the 3-metric. Hence, the Wheeler – DeWitt equation may be interpreted as a hyperbolic differential equation that describes the propagation of wave functions in time hidden in the conformal mode of superspace.

Such an interpretation elevates the canonical Dirac – Wheeler – DeWitt quantization to a new conceptual level similar to the situation at the dawn of the quantum field theory when the substitution of the Shrödinger equation for the relativistic particle by the Klein – Gordon equation has led to second quantization and the description of processes with the variable number of particles. In terms of quantum gravity, such a procedure is usually referred to as third quantization intended to describe creation and annihilation of multiple universes. The necessity of multiple universes in the Klein – Gordon interpretation of the Wheeler – DeWitt equation follows, at least naively, from the properties of the inner product on the space of its solutions.

Due to the hyperbolic nature of this equation and from the analogy with the relativistic particle, the conserved inner product must have the form of a flow across a certain surface Σ in superspace

$$(\boldsymbol{\Psi}_1 | \boldsymbol{\Psi}_2) = \int_{\Sigma} \boldsymbol{\Psi}^{*}_1 {}'' \left\{ \prod_{\mathbf{x}} d\Sigma^{ab}(\mathbf{x}) \times \left(G_{ab,cd} \frac{\hbar \overrightarrow{\delta}}{i\delta g_{cd}(\mathbf{x})} - \frac{\hbar \overleftarrow{\delta}}{i\delta g_{cd}(\mathbf{x})} G_{ab,cd} \right) \right\}'' \boldsymbol{\Psi}_2 . \qquad (2.17)$$

Similar to formula (2.11), the double primes denote here the symbolic character of this expression first proposed by DeWitt in Ref. [6]. The right-hand side of this equation must be constructed on the assumption that this flow is independent of the choice of the surface Σ based on the local conservation law for the quasi-Klein – Gordon current of the DeWitt equation (2.11) and the supermomentum constraints (2.12). However, in quantum cosmology, even the question of dimensionality of Σ is not wholly trivial: this surface does not turn out to be a hypersurface and, hence, the corresponding conservation law is based on Stokes' rather than the Gauss theorem.

It is worth recalling that Eqns (2.11), (2.12) form not a single Klein – Gordon equation but a system of $1 \times \infty^3$ second-order Wheeler – DeWitt equations (one equation per each point of the space) and $3 \times \infty^3$ first order supermomentum equations. Evidently, a set of these $4 \times \infty^3$ equations allows one to formulate the same number of local conservation laws of the continuity equation-type [56]. Hence, the global conservation law based on Stokes' theorem can be formulated for the flow across the surface of dimension $6 \times \infty^3 - 4 \times \infty^3 = 2 \times \infty^3$, by virtue of duality in the superspace of three-dimensional metrics of dimension $6 \times \infty^3$. This quantity is simply a formal dimensionality of the Cauchy surface for the system of Eqns (2.11), (2.12) whereas the surface Σ may be regarded as the initial data surface.

Each of the $1 \times \infty^3$ Wheeler – DeWitt equations generates the local current of the Klein – Gordon type which includes the first-order derivatives of the wave function over superspace coordinates $\overleftrightarrow{\delta}/\delta g_{ab}(\mathbf{x})$. This explains the presence in (2.17) of the product of such factors over points x. Conserved currents of supermomentum equations (2.12) do not contain derivatives with respect to g_{ab} which accounts for their ultralocal contribution in superspace included in the measure $d\Sigma^{ab}$ in the definition of (2.17). The form of this measure and the description of the surface Σ will be given below.

Following these extensive comments on special features of the inner product (2.17) which distinguish it from the simplest relativistic particle case, let us turn to their common properties. The most conspicuous feature shared by the formalisms of the Wheeler – DeWitt and Klein – Gordon equations is that the sign of their conserved currents is undefined: the norm of the real wave function is zero, similar to that of the relativistic particle; it can be negative for complex wave functions due to the presence of derivatives $\overleftrightarrow{\delta}/\delta g_{ab}(\mathbf{x})$ with the Wronskian structure in the measure of the inner product. In particular, for semiclassical wave functions of the form

$$\boldsymbol{\Psi}_{\pm} = P_{\pm} \exp\left(\pm\frac{\mathrm{i}}{\hbar}\boldsymbol{S}\right), \tag{2.18}$$

where S is the Hamilton–Jacobian function satisfying Eqns (2.6)–(2.7), the action of these derivatives leads to the appearance of a factor with the indefinite sign

$$\left(\boldsymbol{\Psi}_{\pm}|\boldsymbol{\Psi}_{\pm}\right) = \int_{\Sigma} |P_{\pm}|^2 \prod_{\mathbf{x}} \mathrm{d}\Sigma^{ab}(\mathbf{x}) \left(\pm G_{ab,\,cd}\,\frac{\delta\boldsymbol{S}}{\delta g_{cd}(\mathbf{x})}\right), \tag{2.19}$$

in the measure, which allows one to consider the two states (2.18) as having norms with opposite signs (hence, having positive and negative frequencies)

$$\left(\boldsymbol{\Psi}_{+}|\boldsymbol{\Psi}_{+}\right) > 0, \qquad \left(\boldsymbol{\Psi}_{-}|\boldsymbol{\Psi}_{-}\right) < 0, \tag{2.20}$$

with respect to such a choice of orientation of $\mathrm{d}\Sigma^{ab}$ at which†

$$\mathrm{d}\Sigma^{ab}(\mathbf{x})\, G_{ab,\,cd}\,\frac{\delta\boldsymbol{S}}{\delta g_{cd}(\mathbf{x})} > 0\,.$$

The absence of a positive-definite inner product rules out the possibility of a meaningful probabilistic interpretation of the wave function $\boldsymbol{\Psi}[g_{ab}, \varphi]$ because natural restriction of the space of solutions of the Wheeler–DeWitt equations to a subset of positive-frequency wave functions results in disregarding the obviously good (from the classical viewpoint) states. Note that in the semiclassical context, the two states (2.18) differ only in that they describe the motion of the system along the same trajectory in superspace in chronologically opposite directions. In contrast to the relativistic particle case, the superspace is *a priori* devoid of any causal structure which makes the two directions of motion equally permissible in terms of the classical approach. On the other hand, superspace has no timelike Killing symmetries (i.e. the supermetric and the potential term of the Wheeler–DeWitt equation are non-trivial functions of superspace coordinates) which might single out preferred positive-frequency solutions. For this reason, the construction of the Hilbert space of states with the positive norm is not unique [57] which provides an additional argument against the restriction of full space of solutions for the Wheeler–DeWitt equations. This accounts for the universally accepted fact that the only possibility is the third quantization, the analog of the transition from quantum mechanics of the relativistic particle to the quantum field theory. Due to this, the key problem, i.e. the indefinite scalar product of cosmological wave functions, is resolved automatically: the wave function of the Universe is no longer the quantum state but becomes an operator acting in a broader Hilbert space of states in which it is possible to introduce a positive-definite inner product. This Hilbert space describes not a single quantum model but many such models, and the operators of solutions with positive frequency for the Wheeler–DeWitt equations induce the creation of new universes whereas negative-frequency operators are responsible for their annihilation.

The idea of third quantization offers conceptual problems, besides purely technical difficulties. The point is that the success of second quantization was essentially due to causality and relativistic invariance of the Klein–Gordon equation theory. Meanwhile, the theory of the Wheeler–DeWitt equation has been mentioned to be *a priori* devoid of both of these properties: the superspace of three-dimensional metrics is curved and although it contains a light cone analog (due to the hyperbolic signature of supermetric‡ (2.16)) this cone does not separate the past and the future. The motion of classical cosmological systems is possible in both directions along the 'timelike' coordinate and outside the cone, in the 'spacelike' direction, the latter being permitted by the variable sign of the potential term in the Einstein–Hamilton–Jacobi equation.

Another problem in the third quantization concept is the construction of an interaction generating quantum transitions with changing number of 'particles', i.e. universes. Direct analogy to the relativistic particle suggests that this construction must be described by the non-linearity of the Wheeler–DeWitt equation, but the approach to this non-linearity should be based on geometrical properties of the processes creating mini- and macrouniverses rather than on the well-known postulates of causality and locality in superspace. In geometrical terms, these processes imply transitions with changing topology of three-dimensional space, that is such geometry of four-dimensional spacetime which makes possible the time-dependent alteration of connectedness of its spatial section (the number of disconnected components or universes). In a typical case, this is geometry of four-dimensional trousers embedded into a certain encompassing

† Strictly speaking, the sign of the infinite product over continuum of points $\prod_{\mathbf{x}}(-1)$ is poorly defined which suggests the incorrectness of formalism in the representation of continuous argument functions. An alternative is the transition to the expansion of $(g_{ab}(\mathbf{x}), \varphi(\mathbf{x}))$ in the discrete basis of some spatial harmonics in which functional superspace arguments are replaced by countable sets of coefficients of such expansion. It seems possible to demonstrate that quantum state frequency is determined by a homogeneous (constant in space) harmonic of such expansion whereas the remaining modes enter in pairs (roughly speaking) and do not contribute to the sign of the state norm. An example is provided by the two-dimensional string model in the basis of paired right ($n > 0$) and left ($n < 0$) moving modes and one spatially homogeneous (zero) mode ($n = 0$). The homogeneous mode in superspace is known to describe the minisuperspace reduction of gravitational systems. Therefore, the frequency of quantum cosmological states turn out to be determined by the minisuperspace sector and, in semiclassical terms, depends on the overall sign of the phase in (2.18). In fact, it is possible to think of an even more complicated situation when the sign of $\mathrm{d}\Sigma^{ab}(\mathbf{x})G_{ab,\,cd}\,\delta\boldsymbol{S}/\delta g_{cd}(\mathbf{x})$ varies in space. However, in such a case, there are points at which this quantity vanishes due to its continuity, and the full measure in the scalar product is nullified by the corresponding factor. Such a situation corresponds to the appearance of a gravitational analog of the problem of Gribov's copies which will be discussed below.

‡ At first sight, the Wheeler–DeWitt equation seems to create an additional problem related to the ultrahyperbolic character of the functional supermetric having signature (2.16) at each point $\mathbf{x}$, hence $1 \times \infty^3$ 'timelike' and $5 \times \infty^3$ 'spacelike' coordinates. However, it should be borne in mind that the Wheeler–DeWitt equation forms ∞^3 equations, instead of one, which are compatible with one another due to the commutation involution with supermomentum equations (see (2.5) and below). This effectively eliminates both multidimensionality of the 'time' coordinate in the hyperbolic differential equation and difficulties encountered in the formulation of the Cauchy problem inalienable in an ultrahyperbolic case.

manifold of large dimension, which describes the collapse of a single closed universe ${}^3\boldsymbol{M}'$ into a pair of universes ${}^3\boldsymbol{M}_1 \cup {}^3\boldsymbol{M}_2$ (Fig. 1) or nucleation of the closed model ${}^3\boldsymbol{M}$ from the open universe ${}^3\boldsymbol{R}'$ (Fig. 2). The interaction in the third quantized theory must be, at least approximately, constructed from the semiclassical amplitude of such processes. It is however known that they lead to the violation of causality in the classical theory with the Lorentzian signature of spacetime [58] and are forbidden by energy considerations in quantization of matter [59]. Therefore, topological transitions may possibly be described by such amplitudes but in a space with the Euclidean signature

$$\exp\left(\mp\frac{1}{\hbar}\,\boldsymbol{I}\right) \tag{2.21}$$

in terms of the Euclidean gravitational action I calculated on the corresponding 4-dimensional spacetime. In the Lagrangian formalism, this action is derived from its Lorentzian analog by analytic continuation to the imaginary time

$$t = -\mathrm{i}\tau, \tag{2.22}$$

which formally makes the metric signature positive-definite. Regions of Euclidean spacetime interpolating between classically allowed Lorentzian 4-geometries (possibly of different topology) constitute the physical picture of semiclassical approximation in the third quantization theory which underlies Sakharov's idea of cosmological transitions with a change of the metric signature [2]. This idea has been realized in the studies of quantum origin of the Universe [17, 18, 60] and in the works exploring the problem of the gravitational loss of quantum coherence in wormhole physics [29, 31, 30].

To summarize, the problem of time in quantum gravity and cosmology gives rise to the third quantization concept which, in turn, leads to the idea of Euclidean-Lorentzian transitions describing effects of gravitational tunnelling in the imaginary time formalism. Possible alternatives to this construction are discussed in the following section.

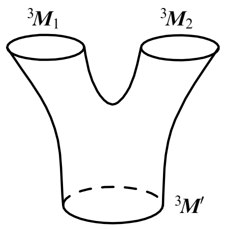

Figure 1. Four-dimensional geometry embedded into a space of larger dimensionality and describing the decay of one closed universe ${}^3\boldsymbol{M}'$ into a pair of universes ${}^3\boldsymbol{M}_1 \cup {}^3\boldsymbol{M}_2$.

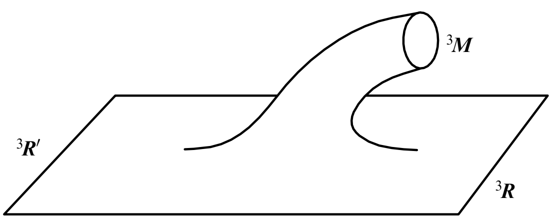

Figure 2. Nucleation of a closed model ${}^3\boldsymbol{M}$ from the open Universe ${}^3\boldsymbol{R}'$.

3. Types of gravity quantization and the state of art: semiclassical methods

The scheme proposed in the previous section just outlines prospects of the closed physical theory whose working mathematical apparatus remains to be constructed. An alternative approach (or approaches) can be based on a more orthodox quantization technique for dynamical systems with constraints that avoids principally new concepts and therefore does not require new physical postulates, ideas of multiple quantum cosmological universes, topological transitions, etc. This section is focused on such methods and will hopefully bring the reader to the belief that third quantization is both necessary and desirable in view of intrinsic difficulties of the formalism which are essentially due to the lack of global applicability of the theory on its full phase space. Restriction to local or quasi-local consideration on either configuration or phase space does not allow one to go beyond the semiclassical expansion which is known to probe only the infinitesimal neighbourhood of the solutions of classical equations. In particular, this explains why we are ready to confine ourselves to loop expansion, the more so that up to now it remains the only method applicable to more or less general realistic problems. It will be shown that such an approach allows for more explicit and rigorous formulation of some problems mentioned above and facilitates their solution.

3.1. The method of semiclassical time

The earliest and seemingly most fruitful (in terms of applicability) method for the introduction of time into quantum cosmology is based on the semiclassical approximation in gravitational field [6, 61, 62, 90]. In this approximation, the wave function is sought in the form

$$\boldsymbol{\Psi}\,[g_{ab},\,\varphi] = \exp\left(\frac{\mathrm{i}}{\hbar}\,\boldsymbol{S}[g_{ab}]\right)|\Psi[g_{ab}]\rangle\,, \tag{3.1}$$

where $\boldsymbol{S}[g_{ab}]$ is the Hamilton-Jacobi function of a purely gravitational field satisfying the vacuum system of Eqns (2.6) and (2.7) without contributions of matter $H_{\perp}^{\mathrm{mat}}$ and H_a^{mat} while the Dirac bra- and ket-notations are used only for quantum states in the Hilbert space of matter fields (the Hilbert space being well-defined and having the positive norm, for it actually represents the conventional quantum field theory in curved space). The substitution of (3.1) into the Wheeler–DeWitt equation yields new equations for the matter field state vector $|\Psi[g_{ab}]\rangle$ parametrically depending on metric

$$''\left\{\frac{\hbar}{\mathrm{i}}\,G_{ab,\,cd}\,\frac{\delta\boldsymbol{S}}{\delta g_{ab}}\,\frac{\delta}{\delta g_{cd}} + \hat{H}_{\perp}^{\mathrm{mat}}(g_{ab}) + \frac{\hbar}{2\mathrm{i}}\,G_{ab,\,cd}\,\frac{\delta^2\boldsymbol{S}}{\delta g_{ab}\delta g_{cd}} - \frac{\hbar^2}{2}\,G_{ab,\,cd}\,\frac{\delta^2}{\delta g_{ab}\delta g_{cd}}\right\}''\,|\Psi[g_{ab}]\rangle = 0, \tag{3.2}$$

$$''\left\{-2\,\frac{\hbar}{\mathrm{i}}\,g_{ab}\nabla_c\,\frac{\delta}{\delta g_{bc}} + \hat{H}_a^{\mathrm{mat}}(g_{ab})\right\}''\,|\Psi[g_{ab}]\rangle = 0. \tag{3.3}$$

In the semiclassical approximation in gravitational field and *for a weak back reaction of quantum matter on the metric background*, these equations can be solved by iterations in powers of the second order terms in functional derivatives, that is, in fact, by the expansion in the small parameter — the ratio of the quantum matter energy density to the Planck one. In the lowest order, it is possible to neglect the third and the fourth terms in (3.2) and consider $|\Psi[g_{ab}]\rangle$ on the solution of classical Einstein's vacuum equations $g_{ab}(\mathbf{x},\, t)$ corresponding to the Hamilton – Jacobi function $\boldsymbol{S}[g_{ab}]$

$$|\Psi(t)\rangle = |\Psi[g_{ab}(\mathbf{x}, t)]\rangle\,. \tag{3.4}$$

With a certain choice of lapse and shift functions ($N^{\perp}$, N^a), this solution satisfies canonical equations

$$\dot{g}_{ab}(t) = N^{\perp} G_{ab,\,cd}\,\frac{\delta \boldsymbol{S}}{\delta g_{cd}} + 2\nabla_{(a}N_{b)}, \tag{3.5}$$

Therefore, the quantum state of matter (3.4) satisfies the evolution equation obtained by substituting (3.5) into

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\,|\Psi(t)\rangle = \int \mathrm{d}^3x \dot{g}_{ab}\,\frac{\delta}{\delta g_{ab}}\,|\Psi[g_{ab}]\rangle \tag{3.6}$$

and taking into account (3.2) – (3.3). The result in the lowest order is the Shrödinger equation for quantized matter fields in the external classical gravitational field which defines the semiclassical time

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\,|\Psi(t)\rangle = \int \mathrm{d}^3x\left\{N^{\perp}\hat{H}^{\mathrm{mat}}_{\perp} + N^a \hat{H}^{\mathrm{mat}}_a\right\}|\Psi(t)\rangle\,. \tag{3.7}$$

Such a method of deriving the quantum field theory from the Wheeler – DeWitt equation dates back, at the model level, to the classical work of DeWitt [6] and was used by Lapchinsky and Rubakov [61] for generic gravitational systems. It turns out to be the most commonly used approach and, in fact, the only currently available convenient way of interpreting the cosmological wave function. It establishes links between fundamental quantum cosmology and applied physics of the early Universe. Physically, this method means that the role of the time variable is played by a gravitational background the quantum properties of which are neglected whereas quantization is entirely carried out in the matter field sector.

Unfortunately,this method is not fundamental despite its significance in applications, because it does not allow the quantum properties of gravitational background to be taken into consideration in a regular manner even in the perturbation theory. The point is that an attempt at solving Eqns (3.2) and (3.3) beyond the lowest order in the quantum-gravitational terms leads to a chain of recurrent equations for quantum corrections which no longer have the form of a homogeneous Shrödinger equation, hence no simple interpretation in terms of unitary evolution is available†. An attempt at consistent unitary quantization of gravity is undertaken in the next section; it is in fact quantization in a reduced phase space first considered in the quantum context by Arnowitt, Deser, and Misner (ADM) [9].

3.2. ADM quantization

Einstein's quantum gravity is actually a special case of the general theory of systems with first class constraints. There are well-established (and equivalent) quantization methods for this theory, at least for the construction of perturbative scattering matrix meeting the unitarity condition [55, 10, 64]. The starting point for all these methods is the reduction of the theory to physical variables in an unitary (canonical) gauge. The reduction procedure as well as subsequent quantization suggests a complicated formalism. To overcome it, one needs condensed DeWitt notations [6] which formally represent the complicated field system in terms of a quantum-mechanical model with the finite dimension of phase space and space of local gauge transformations generated by the constraints. This is achieved by the introduction of such notations for the phase coordinates of the theory

$$q^i = \big(g_{ab}(\mathbf{x}),\, \varphi(\mathbf{x})\big), \qquad p_i = \big(p^{ab}(\mathbf{x}), p_{\varphi}(\mathbf{x})\big), \tag{3.8}$$

in which the condensed index i includes both discrete isotopic indices and the three-dimensional spatial coordinate $\mathbf{x}$. Similar notations for the constraints

$$H_{\mu}(q,\, p) = \big(H_{\perp}(\mathbf{x}),\, H_a(\mathbf{x})\big) \tag{3.9}$$

suggest that the gauge index μ 'labels' the superhamiltonian and supermomenta of the theory as well as their spatial coordinates. Note that functional dependence on field phase-space variables in these notations is represented in the form of functions on phase space (q^i, p_i) while the contraction of the condensed indices includes integration over $\mathbf{x}$, along with discrete summation over spin numbers. In the condensed notations, the canonical action (2.1) has a simple form

$$\boldsymbol{S}[q,\, p,\, N] = \int \mathrm{d}t\,\big\{p_i\,\dot{q}^i - N^{\mu}H_{\mu}(q,p)\big\}, \tag{3.10}$$

and the system of its variational equations

$$\dot{q}^i = \{q^i,\, H_{\mu}\}\,N^{\mu}, \qquad \dot{p}_i = \{p_i, H_{\mu}\}N^{\mu}, \tag{3.11}$$

$$H_{\mu}(q,\, p) = 0 \tag{3.12}$$

conserves constraints in time and leaves the Lagrange multipliers N^{μ}, i.e. lapse and shift functions, absolutely arbitrary since $H_{\mu}(q,\, p)$ are the first class constraints. Their Poisson bracket algebra (2.5) may be written in condensed notations as

$$\{H_{\mu},\, H_{\nu}\} = U^{\alpha}_{\mu\nu}\,H_{\alpha} \tag{3.13}$$

with certain structure functions $U^{\alpha}_{\mu\nu} = U^{\alpha}_{\mu\nu}(q)$. The superhamiltonian and supermomentum constraints $H_{\mu} = (H_{\perp}, H_a)$, quadratic and linear in phase momenta p_i respectively, in these notations have the form

$$H_{\perp} = \frac{1}{2}\,G^{ik}_{\perp}\,p_i\,p_k + V_{\perp}, \qquad H_a = \nabla^i_a\,p_i\,. \tag{3.14}$$

† For example, the third term in Eqn (3.2), $(\hbar/2\mathrm{i})''G_{ab,\,cd}\delta^2\boldsymbol{S}/\delta g_{ab}\delta g_{cd}''$, effectively creates the anti-Hermitian contribution to the Hamiltonian in the Shrödinger equation [63], but its non-unitary interpretation would be incorrect: this term is responsible for the divergence of congruence of classical gravitational trajectories described by the Hamilton – Jacobi function $\boldsymbol{S}$ and can be correctly (in terms of unitarity) taken into account with the help of an one-loop pre-exponential factor of the Pauli – Van Vleck type and a correct measure in the scalar product in the *full* field space. However, such corrections are beyond the scope of the semiclassical time approach and require quantization of *all* field models (see next section).

Here, indices $\perp = (\perp, \mathbf{x})$ and $a = (a, \mathbf{x})$ are also supposed to be condensed, $G_{\perp}^{ik}$ is the ultralocal three-point object containing the DeWitt contravariant supermetric matrix, $V_{\perp}$ denotes the potential term of the superhamiltonian constraint, and ∇_a^i is the generator of infinitesimal spatial diffeomorphism of the variable q^i. $G_{\perp}^{ik}$ and ∇_a^i in the sector of gravitational variables have the form of the following δ-function type kernels:

$$G_{\perp}^{ik} = G_{ab,cd}\delta(\mathbf{x}_i, \mathbf{x}_k)\,\delta(\mathbf{x}_{\perp}, \mathbf{x}_k), \qquad i = (ab, \mathbf{x}_i),$$
$$k = (cd, \mathbf{x}_k), \qquad \perp = (\perp, \mathbf{x}_{\perp}), \tag{3.15}$$

$$\nabla_a^i = -2g_{a(b}\nabla_{c)}\,\delta(\mathbf{x}_a, \mathbf{x}_i),$$
$$i = (bc, \mathbf{x}_i), \qquad a = (a, \mathbf{x}_a), \tag{3.16}$$

which yield local gravitational constraints upon substitution into (3.14) and integration over coordinates†.

Consistent consideration of both quantum and classical dynamics in the theory with first class constraints results in the observation that, first, phase variables of the theory are not independent but obey these constraints (3.12) and, second, they are subject to gauge transformations generated by constraints with the local parameters $\mathcal{F}^{\mu} = \mathcal{F}^{\mu}(t)$

$$\delta q^i = \{q^i, H_{\mu}\}\mathcal{F}^{\mu}, \qquad \delta p_i = \{p_i, H_{\mu}\}\,\mathcal{F}^{\mu}\,. \tag{3.17}$$

This explains why there are much fewer degrees of freedom in the theory than the number of initial pairs of phase variables (q^i, p_i). In order to obtain them, one needs to solve the constraints and separate physical phase variables (ξ^A, π_A) from the purely gauge degrees of freedom in the remaining set of variables. If the initial phase space dimensionality is denoted by $2n$ (in the finite-dimensional context $i = 1, 2, \ldots, n$) and the dimension of the space of constraints as m, $\mu = 1, 2, \ldots, m$, then the dimensionality of physical phase space equals $2(n - m)$, $A = 1, 2, \ldots, n - m$. Technically, phase space variables can be obtained by imposing on $(q^i,\, p_i)$ m canonical additional constraints

$$\chi^{\mu}(q,\, p,\, t) = 0 \tag{3.18}$$

and by solving of the complete system of $2m$ constraint equations and gauge conditions with respect to $2n$ unknown variables $(q^i,\, p_i)$, in terms of physical coordinates and momenta $(\xi^A,\, \pi_A)$. The canonical character of these momenta is guaranteed by the fact that the initial symplectic form $p_i\dot{q}^i$ undergoes transition to the symplectic form on physical phase space $\pi_A\dot{\xi}^A$ plus possible contribution of the Hamiltonian caused by the non-stationary nature of the canonical transformation.

The condition for the separation of physical subspace with the help of gauge (3.18) is the transversality of the gauge surface to the orbit of gauge transformation (3.17). This is guaranteed by non-degeneracy of the Faddeev–Popov operator functional matrix

$$J^{\mu}_{\nu} \equiv \{\chi^{\mu}, H_{\nu}\}, \qquad J = \det J^{\mu}_{\nu} \neq 0\,. \tag{3.19}$$

† It is worthwhile to note that the object $G_{\perp}^{ik}$ itself is not the DeWitt supermetric because it contains an additional δ function absent in the DeWitt supermetric (see Eqn (3.15)). Only functional contraction of $G_{\perp}^{ik}$ and the lapse function $N^{\perp} = 1$, which is constant in space, converts this quantity into the ultralocal metric on superspace $G_{\perp}^{ik}N^{\perp} = G^{ik} = G_{ab,cd}\delta(\mathbf{x}_i, \mathbf{x}_k)$.

The same condition ensures local solvability of the system of constraints (3.12) and gauges (3.18) during reduction to physical variables, as well as the unique choice of Lagrange multipliers, i.e. lapse and shift functions N^{μ} which can be found from the requirement that the gauge conditions be constant in time

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\chi^{\mu} = \{\chi^{\mu}, H_{\nu}\}\,N^{\nu} + \frac{\partial\chi^{\mu}}{\partial t} = 0\,, \tag{3.20}$$

which has only one solution with respect to N^{μ} under the condition (3.19),

$$N^{\mu} = -J^{-1\mu}_{\;\;\;\;\nu}\,\frac{\partial\chi^{\mu}}{\partial t}\,, \tag{3.21}$$

where operator $J^{-1\mu}_{\;\;\;\;\nu}$ is the inverse of J^{μ}_{ν}.

Constraints (3.12) and gauge conditions (3.18) single out $2(n - m)$-dimensional physical subspace in full phase space and, generally speaking, this embedding results in the non-trivial mixing of coordinates and momenta. Such a situation can be substantially simplified by choosing functions of (3.18) to be independent of momenta p_i

$$\chi^{\mu}(q,\, t) = 0, \tag{3.22}$$

which in turn allows physical coordinates ξ^A to be chosen as embedded into n-dimensional superspace of coordinates q^i, with the canonical nature of variables $(\xi^A,\, \pi_A)$ explicitly following from their construction. Indeed, Eqn (3.22) describes embedding a certain $(n - m)$-dimensional submanifold Σ into n-dimensional superspace which is possible to parametrize by arbitrary internal coordinates ξ^A:

$$q^i = e^i(\xi^A,\, t), \qquad \chi^{\mu}\big(e^i(\xi^A, t), t\big) \equiv 0. \tag{3.23}$$

Without the loss of generality, they may be considered as physical configuration coordinates of the theory. It follows from the symplectic form transformation

$$p_i\,\dot{q}^i = p_i\,\frac{\partial e^i(\xi, t)}{\partial\xi^A}\,\dot{\xi}^A + p_i\,\frac{\partial e^i(\xi, t)}{\partial t} \tag{3.24}$$

that the momenta conjugate to ξ^A are equal to the projections of the covector p_i onto the surface Σ

$$\pi_A = p_i\,e^i_A, \qquad e^i_A \equiv \frac{\partial e^i\,(\xi, t)}{\partial\xi^A}\,, \tag{3.25}$$

while the physical Hamiltonian is numerically coincident up to the sign with the second term in (3.24), because the full Lagrangian in the action (3.10) is reduced to the symplectic form $p_i\,\dot{q}^i$ on the solutions of constraints.

Eqns (3.25) define tangential projections of momentum p_i on the surface Σ in terms of physical degrees of freedom. Addition of m constraint equations to these $n - m$ equations yields a system which can be resolved with respect to all momentum components, the solvability condition being again provided by the non-degeneracy of the Faddeev–Popov matrix (3.19). Taken together with (3.23), this completes the procedure of classical reduction to physical degrees of freedom. In physical phase space, the canonical action takes the form

$$S\left[\xi, \pi\right] = \int \mathrm{d}t\,\left[\pi_A\dot{\xi}^A - H_{\mathrm{phys}}(\xi, \pi, t)\right] \tag{3.26}$$

with the physical Hamiltonian

$$H_{\rm phys}(\xi,\pi,t) = -p_i\,(\xi^A,\pi_A,t)\,\frac{\partial e^i(\xi,t)}{\partial t}\,, \qquad (3.27)$$

while all the initial phase variables (q^i, p_i) and the Lagrange multipliers N^μ are constructed as functions of physical coordinates and momenta [8, 56]

$$q^i = e^i(\xi^A,t), \qquad p_i = p_i(\xi^A,\,\pi_A,\,t),$$

$$N^\mu = -J^{-1\mu}_{\;\;\;\;\nu}\,\frac{\partial\chi^\mu}{\partial t}\bigg|_{q^i=e^i(\xi^A,t),\,p_i=p_i(\xi^A,\pi_A,t)}\,. \qquad (3.28)$$

Clearly, the reduction of physical variables results in the solution of the time problem: the action (3.26) acquires the non-zero Hamiltonian by virtue of imposing of the canonical gauge (3.22) explicitly dependent on the time parameter t. The time-dependence of gauge conditions generates the motion of the physical surface $\Sigma(t)$ in superspace, $q^i = e^i(\xi,t)$, and the non-zero lapse and shift functions (3.28). There is a very illuminating interpretation of this phenomenon. The geometrically invariant meaning of the lapse $N^\perp$ and shift N^a functions is in that they are normal and tangential projections of the four-dimensional velocity $\dot X^\alpha(t,\mathbf{x}) \equiv \mathrm{d}X^\alpha(t,\mathbf{x})/\,\mathrm{d}t$ with which the spacelike section $x^\alpha = X^\alpha(t,\mathbf{x})$ propagates in four-dimensional spacetime with coordinates x^α [54]:

$$N^\perp = -n_\alpha \dot X^\alpha(t,\mathbf{x}), \qquad N^a = g^{ab} e^\alpha_b\, {}^4g_{\alpha\beta}\dot X^\beta(t,\mathbf{x})$$

((n_α, e^α_b) forms the basis of the normal and the triad of tangent vectors to hyperspace $t=$ const. This means that the non-stationary surface $\Sigma(t)$ is responsible for the many-fingered time concept in superspace of variables q^i: local variations of this surface parametrized by the parameter t locally generate variations of the lapse and shift functions; hence local variations of constant time hyperspace in four-dimensional space. From the dynamical viewpoint, the introduction of time and the non-zero Hamiltonian is the result of the non-stationary contact canonical transformation from the initial variables $(q^i,\,p_i)$ to physical ones $(\xi^A,\,\pi_A)$. This procedure is inverse of the well-known canonical transformation from current canonical variables to the initial data, constant in time, the trivial dynamics of which is implied by zero Hamiltonian.

Formal quantization of the theory reduced to physical variables encounters no difficulty. Phase-space variables $(\xi^A,\,\pi_A)$ are independent and become, at the quantum level, operators satisfying the Heisenberg commutation relations:

$$[\hat\xi^A,\,\pi_{\hat B}] = \mathrm{i}\hbar\delta^A_B\,.$$

Hermitian realization of the quantum Hamiltonian $\hat H_{\rm phys}$ must be in terms of these Hermitian operators, to define unitary dynamics of physical states $|\Psi(t))$ in the Hilbert space of representation of Heisenberg commutation relations

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\,\big|\Psi(t)\big) = \hat H_{\rm phys}\big|\Psi(t)\big). \qquad (3.29)$$

In the coordinate representation $\big(\xi|\Psi(t)\big) = \Psi(\xi,t)$, $\hat\xi = \xi$, $\hat\pi = \hbar\partial/\mathrm{i}\partial\xi$, the positive-definite scalar product of state vectors with the unit integration measure $\mathrm{d}\xi = \prod_A \mathrm{d}\xi^A$

$$(\Psi_1|\Psi_2) = \int \mathrm{d}\xi\;\Psi_1^*(\xi)\,\Psi_2(\xi) \qquad (3.30)$$

will be the integral of motion of the Shrödinger equation. Then, the knowledge of operator realization of the initial variables of the theory (3.28) with the desired Hermiticity properties will allow one to calculate mean values and matrix elements of all necessary observables of the theory.

ADM quantization in physical space forms a self-consistent closed theory at least up to the problem of operator realization of the physical Hamiltonian and other observables. In the generic non-gauge theory, this problem has no intrinsic solution without an appeal to the experiment. Strictly speaking, there is a similar situation in the theory with constraints. However, in this case, the level of ambiguity in operator realization of the physical Hamiltonian is lower because the gauge theory, even being reduced to the physical sector with certain gauge conditions, retains traces of local gauge invariance explicit in the initial variables (q^i, p_i, N^μ). In the first place, this invariance shows up in a particular dependence of ADM quantization formalism on the choice of gauge conditions. Evidently, the two quantization schemes in the theories obtained by the reduction of one and the same gauge theory in two different gauges must yield physically equivalent results because a change of gauge can be mimicked by gauge transformation. This requirement is trivially fulfilled in the classical approximation and obviously imposes some limitations on the quantum terms in operator realization of the Hamiltonian and other observables. Constructively, this property was realized in gauge theories as a statement that the scattering matrix is independent of the gauge used to build it. However, it was proved in the framework of wider schemes of the BRST (BVF)-type [10] rather than in the language of the reduction to physical sector, largely in terms of formal functional integral which ignores the ordering of local non-commuting operators taken at one point. In the next section, we shall briefly outline how such equivalence of ADM quantum schemes under different gauge conditions can be formulated (at least in the one-loop approximation) as their unitary equivalence to a single Dirac-Wheeler-DeWitt quantization with a special scalar product [65, 8, 56].

3.3 Unitary mapping between the ADM and Dirac-Wheeler-DeWitt quantization schemes.

Note that the physical space section Σ in the previous Section resembles the surface considered in Section 2 and used to construct the conserved inner product in the space of solutions of the Wheeler – DeWitt equations. This observation prompts the idea that various ADM quantization schemes may in fact be projections of one and the same (and independent of the choice of gauges) formalism of the Dirac – Wheeler – DeWitt quantization in superspace. This turns out to be true also for semiclassical states of a certain type in the one-loop approximation.

Going back to the content of Section 2 in the non-condensed DeWitt notations, it should be first of all noted that quantum Dirac constraints on the physical states (2.11), (2.12)

$$\hat H_\mu|\boldsymbol{\Psi}\rangle = 0 \qquad (3.31)$$

must satisfy compatibility conditions which are quantum generalizations of the Poisson bracket algebra (3.13)

$$[\hat{H}_\mu, \hat{H}_\nu] = \mathrm{i}\hbar \hat{U}^\lambda_{\mu\nu} \hat{H}_\lambda \,. \tag{3.32}$$

It is remarkable that these conditions considered as equations for unknown operators $\hat{H}_\mu$, $\hat{U}^\lambda_{\mu\nu}$, which satisfy the correspondence principle with classical theory, may be solved in the subleading approximation in the Planck constant [56]. The answer may be given in the form of a normal qp-ordering $\mathcal{N}_{qp}$ (momenta are to the right of the coordinates) of c-numeric symbols which are expanded in powers of $\hbar$

$$\hat{H}_\mu = \mathcal{N}_{qp}\left\{ H_\mu - \frac{\mathrm{i}\hbar}{2}\frac{\partial^2 H_\mu}{\partial q^i\,\partial p_i} + \frac{\mathrm{i}\hbar}{2}\,U^\nu_{\mu\nu} + O(\hbar^2)\right\}, \tag{3.33}$$

$$\hat{U}^\lambda_{\mu\nu} = \mathcal{N}_{qp}\left\{ U^\lambda_{\mu\nu} - \frac{\mathrm{i}\hbar}{2}\frac{\partial^2 U^\lambda_{\mu\nu}}{\partial q^i\,\partial p_i} - \frac{\mathrm{i}\hbar}{2}\,U^{\lambda\sigma}_{\mu\nu\sigma} + O(\hbar^2)\right\}. \tag{3.34}$$

This result holds for the theories with non-reducible first class constraints of the most general form and includes the higher order structure functions $U^{\lambda\alpha}_{\mu\nu\sigma}$ of gauge algebra [64] which equal zero in the Einstein theory of gravity.

Moreover, in the theory of gravity with constraints of the form (3.14), quadratic and linear in momenta, there is a formal operator realization which closes commutation algebra (3.32) exactly, beyond the perturbation theory in $\hbar$ [66]. It coincides with (3.33), (3.34) in the one-loop approximation and is obtainable from classical gravitational constraints (3.14) by replacing the momenta p_i with functional derivatives $\mathcal{D}_i$ covariant with respect to the Riemannian connection constructed with the help of the DeWitt supermetric and by adding the functional trace of the structural functions $\mathrm{i}\hbar U^\nu_{\mu\nu}/2$ as the anti-Hermitian part:

$$\hat{H}_\perp = -\frac{\hbar^2}{2}\,G^{ik}_\perp \mathcal{D}_i\mathcal{D}_k + V_\perp + \frac{\mathrm{i}\hbar}{2}\,U^\nu_{\perp\nu},$$
$$\hat{H}_a = \frac{\hbar}{\mathrm{i}}\,\nabla^i_a \mathcal{D}_i + \frac{\mathrm{i}\hbar}{2}\,U^\nu_{a\nu}\,. \tag{3.35}$$

It is assumed in the definition of covariant derivatives, that the wave function $\boldsymbol{\Psi}(q)$ at which they act is the *scalar weight of density 1/2*. In the gravitational sector $q^i = g_{ab}(\mathbf{x})$, the contravariant supermetric G^{ik} was defined in the footnote following Eqns (3.15), (3.16); this may be rewritten in condensed notations (and in full metric and matter superspace) in the form of a contraction which converts the three-index object into the two-index one

$$G^{ik} = G^{ik}_\perp N^\perp\big|_{N^\perp=1}\,. \tag{3.36}$$

It has been shown in Refs [8, 66] that, due to the closed algebra of constraints (3.35), this supermetric has generators of spatial diffeomorphisms ∇^i_a

$$\mathcal{D}^i\nabla^k_a + \mathcal{D}^k\,\nabla^i_a = 0, \qquad \mathcal{D}^i = G^{ik}\,\mathcal{D}_k\,, \tag{3.37}$$

as the Killing vectors on superspace. Owing to its ultralocality, $G^{ik} \sim \delta(\mathbf{x}_i, \mathbf{x}_k)$, the covariant derivative preserves not only the metric but also the three-index object $G^{ik}_\perp$: $\mathcal{D}_m G^{ik}_\perp = 0$. Therefore, kinetic terms of operator constraints (3.35) do not require additional ordering prescriptions and, moreover, they are Hermitian relative to the *auxiliary nonphysical* inner product of wave functions (2.14)

$$\langle \boldsymbol{\Psi}_1 | \boldsymbol{\Psi}_2 \rangle = \int \mathrm{d}q\, \boldsymbol{\Psi}^*_1(q)\boldsymbol{\Psi}_2(q)\,. \tag{3.38}$$

However, constraints (3.35) have the anti-Hermitian part as well which is proportional to the trace of structure functions and plays an important role in what follows.

Let us consider now the semiclassical approximation for physical quantum states and look for the solution of the Wheeler–DeWitt equation (3.31) in the form

$$\boldsymbol{\Psi}(q) = \boldsymbol{P}(q)\exp\left[\frac{\mathrm{i}}{\hbar}\,\boldsymbol{S}(q)\right]. \tag{3.39}$$

Here, $\boldsymbol{S}(q)$ is the solution of Hamilton–Jacobi equations which may be briefly represented in condensed notations as

$$H_\mu\left(q, \frac{\partial \boldsymbol{S}}{\partial q}\right) = 0, \tag{3.40}$$

while the expansion of the pre-exponential factor $\boldsymbol{P}(q)$ in powers of $\hbar$ starts with the one-loop contribution which we shall restrict ourselves with. The knowledge of the operator realization of quantum constraints beyond the tree-level approximation allows one to write down the equations for the one-loop prefactor

$$\mathcal{D}_i(\nabla^i_\mu \boldsymbol{P}^2) = U^\lambda_{\mu\lambda}\boldsymbol{P}^2\,, \tag{3.41}$$

$$\nabla^i_\mu \equiv \frac{\partial H_\mu}{\partial p_i}\bigg|_{p\,=\,\partial\boldsymbol{S}/\partial q}\,. \tag{3.42}$$

They have the form of m continuity equations modified by the non-zero right-hand side which originates from the anti-Hermitian contributions to quantum constraints.

It turns out that the solution of these quasi-continuity equations can be written in a closed form [65, 8, 56]. It has the simplest form for the two-point solution of quantum constraints which plays the role of a propagator of physical states in superspace

$$\boldsymbol{K}(q, q') = \boldsymbol{P}(q, q')\exp\left[\frac{\mathrm{i}}{\hbar}\,\boldsymbol{S}(q, q')\right], \tag{3.43}$$

where $\boldsymbol{S}(q, q')$ is also the Hamilton–Jacobi function with respect to both arguments, which represents the classical action (2.8) calculated on the extremal connecting the initial q' and final q points in superspace. The solution of interest is the generalization of the known semiclassical Pauli–Van Vleck determinant of the matrix of second derivatives of this function with respect to end points of the extremal [8]

$$\boldsymbol{S}_{ik'} = \frac{\partial^2 \boldsymbol{S}(q, q')}{\partial q^i\,\partial q^{k'}}\,. \tag{3.44}$$

The determinant of this matrix is zero because of its degeneracy caused by zero-vectors (3.42)

$$\nabla^i_\mu \boldsymbol{S}_{ik'} = 0, \tag{3.45}$$

following from the differentiation of Eqns (3.40) with respect to q' (and by analogous right-hand zero-vectors $\nabla^{k'}_\nu$). However, there is an invariant procedure for calculating this determinant on the non-degeneracy subspace of the matrix (3.44), which is equivalent to the one-loop gauge fixing procedure of Faddeev and Popov, leading to the solution of Eqn (3.41). This procedure consists in the introduction of arbitrary covariant vectors χ^μ_i at the point q having the non-degenerate matrix of scalar products with zero-vectors ∇^i_μ and analogous vectors $\chi^\nu_{k'}$ at the point q'

$$J^\mu_\nu = \chi^\mu_i \nabla^i_\nu, \qquad J \equiv \det J^\mu_\nu \neq 0,$$
$$J'^\mu_\nu = \chi^\mu_{i'} \nabla^{i'}_\nu, \qquad J' \equiv \det J'^\mu_\nu \neq 0. \tag{3.46}$$

Using the additional reversible matrix $c_{\mu\nu}$, these vectors allow one to replace the zero-blocks of the matrix $\boldsymbol{S}_{ik'}$ by the gauge-fixing term, converting it into the reversible operator

$$\boldsymbol{F}_{ik'} = \boldsymbol{S}_{ik'} + \chi^\mu_i c_{\mu\nu} \chi^\nu_{k'}, \tag{3.47}$$

and to construct the one-loop pre-exponential factor in terms of its own and 'Faddeev – Popov' determinants (3.46)

$$\boldsymbol{P} = \left[\frac{\det \boldsymbol{F}_{ik'}}{JJ' \det c_{\mu\nu}}\right]^{1/2}. \tag{3.48}$$

This expression does not depend on the choice of auxiliary objects $(\chi^\mu_i, \chi^\nu_{k'}, c_{\mu\nu})$ which fix the gauge, due to a certain analog of the Ward identity [8, 56], and is the solution of the 'continuity' equations (3.41).

Let us consider the projection of the constructed propagator (3.43) with the pre-exponential factor (3.48) onto the physical subspace (3.23), in order to establish the promised unitary map between the Dirac – Wheeler – DeWitt and ADM quantization schemes. Geometry of the local embedding of Σ into superspace of coordinates q^i has been considered in much detail in Ref. [8]. Here, it is worth noting that this embedding can conveniently be described in special coordinates on superspace $\bar{q}^i = (\xi^A, \theta^\mu)$ in which ξ^A are internal coordinates on Σ (ADM physical configuration coordinates), and θ^μ are defined by equations of gauge conditions

$$q^i \to \bar{q}^i = (\xi^A, \theta^\mu), \quad q^i = e^i(\xi^A, \theta^\mu), \quad \theta^\mu = \chi^\mu(q, t). \tag{3.49}$$

The equation of the surface in the new coordinates is trivial $\theta^\mu = 0$, while the equation of embedding (3.23) coincides with this reparametrization equation at $\theta^\mu = 0$, $e^i(\xi, t) = e^i(\xi, 0, t)$. Following from this reparametrization is the relation between integration measures on superspace $\mathrm{d}q = \mathrm{d}^n q$ and on Σ, $\mathrm{d}\xi = \mathrm{d}^{n-m}\xi$

$$\mathrm{d}\xi = \mathrm{d}q\, \delta(\chi)\, M, \quad \delta(\chi) = \prod_\mu \delta\big(\chi^\mu(q, t)\big),$$
$$M = \big(\det\,[e^i_A, e^i_\mu]\big)^{-1}, \tag{3.50}$$

where the transformation Jacobian is built in terms of the basis of vectors tangential and normal to the surface Σ

$$e^i_A = \frac{\partial e^i}{\partial \xi^A}, \qquad e^i_\mu = \frac{\partial e^i}{\partial \theta^\mu}.$$

Note that m covariant vectors normal to Σ are written as gauge gradients

$$\chi^\mu_i = \frac{\partial \chi^\mu}{\partial q^i}, \qquad \chi^\mu_i e^i_\nu = \delta^\mu_\nu, \tag{3.51}$$

with which it is possible (without the loss of generality) to identify auxiliary covectors used in the construction of the one-loop pre-exponential factor above. It is also possible to choose a combination of vectors e^i_A tangent to Σ and vectors (3.42) as a complete local basis, due to the non-degeneracy of the Faddeev – Popov operator (3.19) which coincides in this case with (3.46). Vectors (3.42) are re-expanded in basis (e^i_A, e^i_μ) in accordance with $\nabla^i_\mu = e^i_\nu J^\nu_\mu + e^i_A (\ldots)^A$, whence the following relation for the matrix determinant of the new basis vectors (e^i_A, ∇^i_μ) holds

$$\det\,[e^i_A, \nabla^i_\mu] = \frac{J}{M}. \tag{3.52}$$

The projection of the propagator kernel (3.43) with respect to its two arguments q and q' on the two corresponding physical spaces $\Sigma(t)$ and $\Sigma(t')$ is performed using the explicit relation between the Hamilton – Jacobi functions on superspace $\boldsymbol{S}(q, q')$ and on the space of ADM physical variables

$$S\big(t, \xi | t', \xi'\big) = \boldsymbol{S}\big(e\,(\xi, t),\ e\,(\xi', t')\big) \tag{3.53}$$

and also by expanding the matrix $\boldsymbol{F}_{ik'}$ in the basis (e^i_A, ∇^i_μ)†. The result of the projection

$$K\big(t, \xi | t', \xi'\big) = \mathrm{const}\left(\frac{J}{M}\right)^{1/2} \times \boldsymbol{K}(q, q') \left(\frac{J'}{M'}\right)^{1/2} \Bigg|_{q = e(\xi, t),\ q' = e(\xi', t')} \tag{3.54}$$

is the one-loop unitary propagator of the ADM quantum scheme given by the known Pauli – Van Vleck formula [67]

$$K\big(t, \xi | t', \xi'\big) \equiv \left[\det \frac{\mathrm{i}}{2\pi\hbar} \frac{\partial^2 S}{\partial \xi^A \partial \xi^{B'}}\right]^{1/2} \exp\left[\frac{\mathrm{i}}{\hbar} S\big(t, \xi | t', \xi'\big)\right]. \tag{3.55}$$

It satisfies the Shrödinger equation with the Hermitian operator realization of the physical Hamiltonian (3.27)‡.

The relation (3.54) is the desired unitary map between the Dirac – Wheeler – DeWitt quantization in superspace and ADM quantization in physical variables. This map for propagators obviously implies the map for wave functions $\boldsymbol{\Psi}(q)$ and $\Psi(t, \xi)$ of these two schemes

$$\Psi(t, \xi) = \left(\frac{J}{M}\right)^{1/2} \boldsymbol{\Psi}(q) \Bigg|_{q = e(\xi, t)}, \tag{3.56}$$

and is actually unitary provided that the physical scalar products of these two states coincide in both schemes

$$(\Psi | \Psi) = (\boldsymbol{\Psi} | \boldsymbol{\Psi}). \tag{3.57}$$

Proceeding from the ADM scalar product (3.30) and substituting (3.56) into the left-hand side of this equation, one obtains, due to the replacement of integration variables (3.50), an expression for the *physical* inner product of the

† The Hamilton – Jacobi function of the system reduced to ADM physical variables explicitly coincides with the action functional on the extremal which connects points ξ and ξ' at moments t and t' and satisfies the conventional Hamilton – Jacobi equation $\partial S/\partial t + H_{\rm phys}(\xi, \partial S/\partial \xi) = 0$ with the non-zero physical Hamiltonian (3.27).

‡ This realization is analogous to (3.33), (3.34) and satisfies the principle of correspondence with classical theory [8, 56]: $\hat{H}_{\rm phys} = \mathcal{N}_{\xi\pi}[\,H_{\rm phys} - (\mathrm{i}\hbar/2)\,\partial^2 H_{\rm phys}/\partial\xi^A \partial\pi_A + O(\hbar^2)\,]$.

Dirac wave functions

$$(\boldsymbol{\Psi}\,|\,\boldsymbol{\Psi}) = \int \mathrm{d}q\, \delta\big(\chi(q,t)\big)\, \boldsymbol{\Psi}^*(q)\, J\, \boldsymbol{\Psi}(q). \tag{3.58}$$

Eqns (3.55) – (3.58) constitute the essence of the unitary reduction from the Dirac – Wheeler – DeWitt quantization to the ADM quantization scheme. The ADM quantum schemes in different gauges (hence, at different choices of physical time and/or at different moments of the one and the same time) can be obtained by projecting the superspace wave function $\boldsymbol{\Psi}(q)$, independent of gauges and free of any information about them, onto different physical subspaces Σ. This picture is consistent largely due to time and gauge-independence of the physical inner product (3.58). Unitarity of the ADM quantum scheme accounts for t-independence of this inner product, hence the integral on the right-hand side of (3.58) actually taken over the surface $\Sigma(t)$ in superspace must not depend on the choice of this surface. Evidently, in the Dirac-Wheeler – DeWitt quantization formalism, this property must be the result of quantum constraints (3.31) on wave functions which does not depend on a given ADM reduction to the physical sector. This is exactly the case in the one-loop approximation considered here.

In order to demonstrate this, let us note that the measure of integration over $\Sigma(t)$ in (3.58) is non-trivial and, for semiclassical states of the form (3.39) in the one-loop approximation, is given by the state-dependent Faddeev – Popov determinant

$$J = \det\left[\frac{\partial\chi^\mu(q,t)}{\partial q^i}\,\frac{\partial H_\nu(q,p)}{\partial p_i}\right]_{p\,=\,\partial \boldsymbol{S}/\partial q} = \det\,(\chi^\mu_i \nabla^i_\nu). \tag{3.59}$$

Using the fact that $\mathrm{d}q^i = e^i_A\,\mathrm{d}\xi^A$ on Σ and Eqns (3.50) and (3.52), it is easy to rewrite the inner product (3.58) as the integral of the external $(n-m)$-form over the oriented $(n-m)$-dimensional surface Σ

$$(\boldsymbol{\Psi}\,|\,\boldsymbol{\Psi}) = \int_\Sigma \omega^{(n-m)}, \tag{3.60}$$

$$\omega^{(n-m)} = \frac{\mathrm{d}q^{i_1}\wedge\ldots\wedge \mathrm{d}q^{i_{n-m}}}{m!(n-m)!} \times \epsilon_{i_1\ldots i_n}\boldsymbol{\Psi}^*(q)\,\nabla^{i_{n-m+1}}_{\mu_1}\ldots\nabla^{i_n}_{\mu_m}\,\boldsymbol{\Psi}(q)\,\epsilon^{\mu_1\ldots\mu_m}, \tag{3.61}$$

where $\epsilon_{i_1\ldots i_n} = \pm 1$ and $\epsilon^{\mu_1\ldots\mu_m} = \pm 1$ are completely antisymmetric tensor densities on n- and m-dimensional spaces, respectively†. Using the continuity equations (3.41) for the pre-exponential factor of wave functions and the closed Lie brackets algebra of vector fields ∇^i_μ (a consequence of the closed algebra of constraints [8, 56])

$$\nabla^i_\mu \mathcal{D}_i \nabla^n_\nu - \nabla^i_\nu \mathcal{D}_i \nabla^n_\mu = U^\alpha_{\nu\mu}\nabla^n_\alpha, \tag{3.62}$$

it is equally easy to demonstrate that this form is closed

$$\mathrm{d}\omega^{(n-m)} = 0. \tag{3.63}$$

Hence, by virtue of the Stokes' theorem for a $(n-m+1)$-dimensional surface $\mathcal{D}$ responsible for cobordism of two physical (oppositely oriented) spaces Σ and Σ' ($\partial\mathcal{D} = \Sigma\cup\Sigma'$),

$$\int_\Sigma \omega^{(n-m)} - \int_{\Sigma'} \omega^{(n-m)} = \int_{\mathcal{D}} \mathrm{d}\omega^{(n-m)} = 0, \tag{3.64}$$

it follows that the inner product is independent of the choice of Σ. This simple property accumulates both unitarity and gauge-independence of the Dirac – Wheeler – DeWitt quantization formalism.

This integral of motion of the Wheeler – DeWitt equations is nothing but the semiclassical conservation law for their current discussed in Section 2. Comparison of (3.60) and (3.61) with (2.17) and (2.19) is based on the fact that in the exterior form (3.61) from the alternated product ∇^i_μ, $\nabla^i_\mu = (\nabla^i_\perp, \nabla^i_a)$, it is possible to single out as a factor only the product of quantities $\nabla^i_\perp$. If one assumes that the condensed index $\perp = \mathbf{x}$ can take values $\perp = 1, 2, \ldots, M$ (actually, $M = \infty^3$), then this $(n-m)$-form becomes

$$\omega^{(n-m)} = \nabla^{i_1}_1\ldots\nabla^{i_M}_M\,\mathrm{d}^{n-m}\Sigma_{i_1\ldots i_M}, \tag{3.65}$$

where the measure $\mathrm{d}^{n-m}\Sigma_{i_1\ldots i_M}$ includes all the rest (including all gauge-fixing elements for spatial diffeomorphisms). Bearing in mind that $\nabla^i_\perp = G^{ik}_\perp\times\times\partial\boldsymbol{S}/\partial q^k$, it is possible to arrive, in the linear in $\hbar$ approximation, at both representations (2.17) and (2.19) for the scalar product of semiclassical wave functions (3.39) considered in Section 2‡

$$(\boldsymbol{\Psi}\,|\,\boldsymbol{\Psi}) = \int_\Sigma \big|P(q)\big|^2\left(\prod_\perp G^{i_\perp k}_\perp\,\frac{\partial\boldsymbol{S}}{\partial q^k}\right)\mathrm{d}^{n-m}\Sigma_{i_1\ldots i_M}$$
$$= \int_\Sigma \boldsymbol{\Psi}^*(q)\,\prod_\perp\frac{\hbar}{2\mathrm{i}}\left(G^{i_\perp k}_\perp\,\frac{\overrightarrow{\partial}}{\partial q^k} - \frac{\overleftarrow{\partial}}{\partial q^k}\,G^{k i_\perp}_\perp\right)$$
$$\times\,\boldsymbol{\Psi}(q)\,\mathrm{d}^{n-m}\Sigma_{i_1\ldots i_M}\,. \tag{3.66}$$

Thus, we have shown that it is possible to come to the Dirac – Wheeler – DeWitt quantization scheme in superspace of 3-metrics and matter fields, based on quantization in physical variables obtained by the ADM reduction. The physical inner product of ADM wave functions (with the trivial measure) coincides, at least semiclassically, with the conserved current of the Wheeler – DeWitt equations of the quasi-Klein-Gordon type. Because of the complicated non-Abelian nature of the gravitational field, it is known only in the one-loop approximation, similar to quantum reduction at large, from the Dirac – Wheeler – DeWitt to ADM formalism. Obviously, the entire operator scheme of unitary equivalence can be qeneralized to the highest loop-orders. This statement is supported by powerful theorems of gauge-independence of the formal functional integral in generic gauge theories [10, 64]§. In this case, however, one encounters the same

† In the context of infinite-dimensional configuration ($n = 6\times\infty^3$) and gauge ($m = 4\times\infty^3$) spaces, these tensor densities are certainly formal, and the factorial coefficients should be understood as certain infinite normalization multipliers.

‡ Strictly speaking, the complete agreement with the heuristic formulas (2.17) and (2.19) is hardly possible since the factorization of the measure over space points in the form of $\mathrm{d}^{n-m}\Sigma_{i_1\ldots i_M} = \prod_\perp \mathrm{d}\Sigma_{i_\perp}$ is either unattainable or exists in a narrow class of special gauge conditions.

§ Note that the integration measure in both the gauge field functional integral and the expression for the inner product (3.58) equally includes the gauge condition and the Faddeev-Popov determinant, the only essential difference being the functional dimensionality of the integration space and the determinant: the space of time histories versus the space of configuration coordinates at a fixed moment. Therefore, the nature of gauge-independence as the independence of the choice of physical space embedded into a wider gauge field space is the same in both cases and requires similar methods.

difficulties as in Section 2 which will be shown to serve as a motivation for third quantization.

3.4. The problem of Gribov's copies, origin of Euclidean regions, and motivation for third quantization

The ADM reduction to physical variables and the interpretation of the Dirac–Wheeler–DeWitt formalism in terms of averages (3.58) are essentially based on non-degeneracy of the Faddeev–Popov operator as a function on phase space (3.19) or a function on superspace (3.59) in the semiclassical Hamilton–Jacobi theory. For superspace gauges of the form (3.22), the 'column' of this matrix $J^{\mu}_{\perp}$ is linear in momenta

$$J^{\mu}_{\perp}(q,p) = \frac{\partial \chi^{\mu}(q,t)}{\partial q^i}\, G^{ik}_{\perp} p_k\,, \tag{3.67}$$

and certainly leads to its degeneracy on a certain submanifold which includes $p_i = 0$. The determinant $J = \det J^{\mu}_{\nu}$ vanishes on this submanifold and changes sign when crossing it. This means that globally, on phase space in this type of gauge, the ADM reduction procedure does not work: the solution for constraint equations with respect to momenta p_i in (3.28) is no longer unique and the corresponding lapse and shift functions become singular. Such a property is a canonical analogue of the problem of Gribov's copies [68] when the system of constraint and gauge equations does not single out a single representative from the class of gauge-equivalent configurations in the phase space. In the gravitational context, where gauge imposition is concurrently used to introduce time, this property means that the introduced variable does not cover globally and unambiguously dynamical evolution of physical degrees of freedom, that is it is not a function monotonically changing along classical (and quantum) histories†.

At the quantum level in the one-loop approximation, the same property means that the sign of the inner product (3.58) is indefinite which leads to the third quantization concept, as was discussed in Section 2. This concept is known to be related to Euclidean–Lorentzian transitions between states with different spacetime signatures. Our objective is to show how transitions of this type actually arise from global non-positiveness of the Faddeev–Popov determinant, when submanifolds of its zeros describe sections in spacetime on which a change of metric signature occurs.

It is well-known from quantum mechanics and the theory of gauge field instantons that the imaginary time formalism (2.22) (or the formalism of Euclidean spacetime) is used to describe underbarrier effects of quantum tunnelling related to the location of the system in classically forbidden regions of the phase space. For quantum states of the form (3.39), a classically forbidden region in superspace arises when the congruence of extremals of classical equations defined by the function $\boldsymbol{S}(q)$ does not reach points of this region and is reflected from the caustic which separates it from the classically allowed part of superspace. In the context of one-dimensional configuration space, this is simply a turning point whereas in multidimensional problems, the caustic is a submanifold which may be described in the following way. Let

† Generally speaking, the sign of the lapse function in (3.28) changes when the determinant $J = \det J^{\mu}_{\nu}$ crosses zero which means an actual decrease in physical time of the observer $N^{\perp}\,\mathrm{d}t < 0$ as the parameter t continues to monotonically grow. Transition from the Lorentzian to Euclidean regime corresponds to a change of the sign in $N^{\perp 2}$, i.e. complexification of $J^{\mu}_{\perp}$.

$$q^i = q^i(t,\xi') \tag{3.68}$$

represent a flow of classical trajectories in superspace described by the function $\boldsymbol{S}(q)$ and starting at $t = 0$ at the points $\xi' = \xi^{A'}$ of the initial Cauchy surface (ADM physical space at the initial moment). The initial physical coordinates label individual trajectories in their bunch (3.68), and derivatives with respect to them form the system of vectors $q^i_{A'} = \partial q^i(t,\xi')/\partial\xi^{A'}$ transversal to the bunch. By virtue of canonical equations of motion, tangent vectors to the trajectories are equal to

$$\dot{q}^i(t,\xi') = N^{\mu}\, \frac{\partial H_{\mu}}{\partial p_i}\bigg|_{p = \partial \boldsymbol{S}/\partial q} = N^{\mu}\nabla^i_{\mu}\,. \tag{3.69}$$

The caustic emerges when the congruence of trajectories gets folded in such a way that their tangent vector falls in the plane spanned by transversal vectors $q^i_{A'}, \dot{q}^i(t,\xi) = v^A q^i_{A'}$ which means a linear dependence of $q^i_{A'}$ and ∇^i_{μ}:

$$\det\,[q^i_{A'}, \nabla^i_{\mu}] = 0. \tag{3.70}$$

On the other hand, the embedding into a superspace of physical space $\Sigma(t)$, to which the point (3.68) belongs, implies that

$$q^i(t,\xi') = e^i\big(\xi(t,\xi'),t\big),$$
$$q^i_{A'} = e^i_B u^B_{A'}, \qquad u^B_{A'} \equiv \frac{\partial \xi^B(t,\xi')}{\partial \xi^{A'}}\,, \tag{3.71}$$

where $\xi(t,\xi')$ is a congruence of classical trajectories in the ADM physical space parametrized by the initial data ξ', while the reversible matrix $u^B_{A'} = u^B_{A'}(t)$ forms the system of basis functions of linearized equations of motion in the reduced theory. Therefore, the equation for the envelope of the family of trajectories (3.70) takes (in view of (3.52)) the form

$$\det\left[e^i_{A'}, \nabla^i_{\mu}\right] \det u^A_{B'} = \det u^A_{B'}\; \frac{J}{M} = 0. \tag{3.72}$$

It is not difficult to choose the physical embedding in superspace in such a way that the Jacobian M in (3.50) is globally non-degenerate. Therefore, the condition of the caustic in *superspace* breaks into two: the condition for the formation of the caustic in *physical* configuration space

$$\det u^A_{B'}(t) = 0, \tag{3.73}$$

and/or the condition for the formation of Gribov's copies

$$J = 0. \tag{3.74}$$

Unlike a theory without constraints, quantum gravitation has freedom conditioned by the difference between caustics in physical space Σ and superspace. Analysis of specific cases of *multi-dimensional* gravitational tunnelling, e.g. in a model of quantum creation of a chaotic inflationary Universe (see Section 4), shows that there are caustics in superspace without singularities (3.73), although Eqn (3.74) is satisfied‡

‡ Note that the problem (3.73) can be eliminated by the identification of physical variables and initial data ξ' which implies $u^A_{B'}(t) = \delta^A_{B'}$ and converts $\det u^A_{B'}(t)$ into unity. This means that Eqn (3.68) is considered to be a function of physical space embedding into superspace. Transition from $\xi(t)$ to ξ' is evidently a non-stationary canonical transformation in ADM variables which again nullifies the physical Hamiltonian obtained by the ADM reduction from the parametrized theory with the Hamiltonian vanishing on gravitational constraints.

and therefore has an invariant content and is inseparable from the problem of Gribov's copies. Here, Gribov's copies at the qualitative level appear when physical space crosses the classical extremal at least twice: one branch is crossed before the caustic and the other — after the reflection from the caustic. Hence, ambiguity of the ADM reduction and the indefinite sign of the physical state norm arise due to the contribution of opposite signs originating from these branches.

It is well-known that the extension of trajectories beyond the caustic is possible at the cost of complexification of the time argument in the solutions of equations (the simplest case being the Wick rotation (2.22)) and, in the general case, by the complexification of configuration variables themselves. The former case has been intensively considered in the context of under-barrier phenomena described by instantons whereas complex tunnelling was given less attention because it is exponentially suppressed in conventional theories of the non-gravitational type [69]. In the gravitational context, 'real' tunnelling known as Euclidean quantum gravity was also extensively studied [70, 71] and served as a basis of the wormhole physics [30]. However, the theory of gravitational tunnelling is wider than that of Ref. [69], and one of the most interesting applications of Euclidean quantum gravity, the theory of the quantum origin of the inflationary Universe, deals (beyond the leading approximation) with complex fields and geometries. We shall not consider complex tunnelling here (elements of its general theory were discussed in Ref. [69]) and restrict ourselves with a brief summary of the present section: that there is a class of gravitational Gribov's copies connected with caustics in the Einstein–Hamilton–Jacobi theory whose equation is given by zeroes of the Faddeev-Popov determinant. The corresponding points in superspace are mapped on the spacetime section; at these points, a change of the signature (in the case of real tunnelling) and complexification of classical extremals themselves (in a more complicated case) occur. In either case, the Hamilton–Jacobi function acquires the imaginary part

$$\boldsymbol{S} \to \boldsymbol{S} + \mathrm{i}\boldsymbol{I}\,, \tag{3.75}$$

which, in the former case, can be identified with the Euclidean gravitational action calculated on regions with the positive signature and defining the semiclassical amplitude (2.21). This amplitude characterizes (non-perturbative in $\hbar$) effects of the Universe creation which can probably be consistently interpreted only in terms of third quantization. It is therefore clear that, proceeding from a consistent gravity quantization scheme by means of preliminary reduction to the ADM physical sector, we have to deal with the problem of Gribov's copies which once again brings us to the problem of third quantization.

Now, we can tell much more about the formalism of third quantization than in Section 2. In particular, we can see, regardless of any model restrictions, that in the one-loop approximation for 'positive'-frequency ($J > 0$) solutions of the form (3.39), the inner product of ADM quantization coincides with the conserved 'quasi-Klein-Gordon' current of the Wheeler–DeWitt equations (3.58) or (3.66). However, this current is known only in the one-loop approximation and, for the general semiclassical states of the form

$$\boldsymbol{\Psi}(q) = \sum_I \boldsymbol{P}_I(q) \exp\left[\frac{\mathrm{i}}{\hbar}\,\boldsymbol{S}_I(q)\right], \tag{3.76}$$

including negative-frequency components, it is conserved only in the form (3.66) where momenta in the measure $J(q, \partial\boldsymbol{S}/\partial q)$ are replaced by Wronskian type operators

$$p = \frac{\partial \boldsymbol{S}}{\partial q} \to \frac{\hbar}{2\mathrm{i}}\,\overleftrightarrow{\partial}_q = \frac{\hbar}{2\mathrm{i}}\left(\frac{\overrightarrow{\partial}}{\partial q} - \frac{\overleftarrow{\partial}}{\partial q}\right), \tag{3.77}$$

that act on the wave functions $\boldsymbol{\Psi}(q)$ and $\boldsymbol{\Psi}^*(q)$, which may be rewritten with the same accuracy as†

$$(\boldsymbol{\Psi}\,|\,\boldsymbol{\Psi}) = \int \mathrm{d}q\,\delta\big(\chi(q,t)\big)\,\boldsymbol{\Psi}^*(q)\,J\left(q, \frac{\hbar}{2\mathrm{i}}\,\overleftrightarrow{\partial}_q\right) \times \boldsymbol{\Psi}(q) + O(\hbar)\,. \tag{3.78}$$

Given negative frequencies in the superposition (3.76), unitary mapping (3.56) between the ADM quantum states and Dirac–Wheeler–DeWitt states does not make sense even in the one-loop approximation. This is quite evident from the factor $(J/M)^{1/2}$ in (3.55) which is imaginary for the components of the wave function having a negative frequency. Therefore, the interpretation of the Wheeler–DeWitt equations in terms of unitary reduction to ADM quantization may be correct only for a subclass of their solutions ($J > 0$), or may be an approximate concept disregarding the interaction between the sectors of opposite frequencies.

3.5. Relativistic particle and second quantization

A relativistic particle provides a simple example when the quantum-mechanical problem of Gribov's copies serves as a motivation for second quantization. Dynamics of the relativistic particle follows from the action

$$\boldsymbol{S}\big[x(t)\big] = -m\int \mathrm{d}t\big(-g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta\big)^{1/2},$$
$$g_{\alpha\beta} = \mathrm{diag}\,(-1,1,1,1), \qquad \alpha = 0,1,2,3, \tag{3.79}$$

which gives rise, in the canonical formalism, to the superhamiltonian constraint on the phase space of coordinates $q^i = x^\alpha$ and momenta $p_i = p_\alpha$

$$H \equiv g^{\alpha\beta} p_\alpha p_\beta + m^2 = 0\,, \tag{3.80}$$

which, in the Dirac quantization scheme, becomes the Klein–Gordon equation for the wave function:

$$(-\hbar^2\Box + m^2)\,\boldsymbol{\Psi}(x) = 0. \tag{3.81}$$

† Conservation of diagonal terms is due to the mechanism discussed above. For non-diagonal terms (integrals of the products of the I-th and J-th components with $I \neq J$ in (3.76)), the steepest descent method equates only the projections of the I-th and J-th momenta tangential to Σ at the stationary point (since the integral is taken only over Σ) while the normal components from the solution of constraint equations have the same absolute values. However, the action of the Wronskian leads to the sum of these normal components in the measure; therefore they must have the same sign to avoid cancellation. This property accounts for the necessity of the Wronskian-type operator measure to ensure the orthogonality of components with different frequencies and the conservation of the norm of superpositions (3.76). Further proof that non-diagonal contributions with $\boldsymbol{S}_I(q) \neq \boldsymbol{S}_J(q)$ vanish perturbatively is standard: there can be no identical real extremals for different Hamilton–Jacobi functions satisfying the same Cauchy data, therefore the saddle point must lie in the complex plane and its contribution must be exponentially suppressed by the imaginary part of the action. A similar proof of conservation (perturbative suppression) holds for matrix elements of different quantum states.

It is easy to demonstrate that the algorithm (3.43) – (3.48) for the two-point solutions of this equation with two Hamilton – Jacobi functions of the form

$$\boldsymbol{S}(x, x') = \mp m\sqrt{-g_{\alpha\beta}(x^\alpha - x^{\alpha\prime})(x^\beta - x^{\beta\prime})} \tag{3.82}$$

leads to Wightman functions having positive and negative frequencies respectively [8, 56]. The ADM reduction in the natural gauge

$$\chi(x, t) \equiv x^0 - t = 0 \tag{3.83}$$

identifies physical degrees of freedom with spatial components of coordinates and momenta $(\xi^A, \pi_A) = (\mathbf{x}, \mathbf{p})$ and is ambiguous: the two physical Hamiltonians are in fact two roots of the superhamiltonian constraint

$$H_{\text{phys}}(\mathbf{x}, \mathbf{p}) = \pm\omega_{\mathbf{p}}, \quad \omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}\,. \tag{3.84}$$

This ambiguity means the presence of two Gribov's copies separated by zeroes of the Faddeev – Popov 'determinant'

$$J = 2p^0. \tag{3.85}$$

In the absence of external fields for a massive particle, these zeroes lie in a classically forbidden region $\mathbf{p}^2 + m^2 = 0$; therefore, the two sectors of solutions having positive and negative frequencies are explicitly separated. Solutions with positive frequency, specifically

$$\boldsymbol{\Psi}_{(+)}(x) = \int \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}}\, a(\mathbf{p})\, \exp\left[\mathrm{i}(-\omega_{\mathbf{p}} x^0 + \mathbf{p}\cdot\mathbf{x})\right], \tag{3.86}$$

are unitarily mapped onto physical Newton – Wigner wave functions

$$\Psi_{(+)}(x) = \int \frac{\mathrm{d}^3 p}{(2\pi\hbar)^{3/2}}\, a(\mathbf{p})\, \sqrt{2\omega_{\mathbf{p}}}\, \exp\left[\mathrm{i}(-\omega_{\mathbf{p}} x^0 + \mathbf{p}\cdot\mathbf{x})\right], \tag{3.87}$$

where $\sqrt{2\omega_{\mathbf{p}}}$ is the factor $(J/M)^{1/2}$, $M = 1$, of Eqn (3.55) in the momentum representation. The inner product on the space of solutions for the Klein – Gordon equation for positive frequencies is positive and coincides with the inner product of the Newton – Wigner states

$$\begin{aligned}\left(\widetilde{\boldsymbol{\Psi}}_{(+)}|\boldsymbol{\Psi}_{(+)}\right) &= -\frac{\hbar}{2\mathrm{i}}\int \mathrm{d}^4 x\, \delta(x^0 - t)\, \widetilde{\boldsymbol{\Psi}}^*_{(+)}(x)\, \overleftrightarrow{\partial}_0 \boldsymbol{\Psi}_{(+)}(x) \\ &= \int \mathrm{d}^3 p\, (2\omega_{\mathbf{p}})\, \widetilde{a}^*(\mathbf{p}) a(\mathbf{p}) = \left(\widetilde{\Psi}|\Psi\right).\end{aligned} \tag{3.88}$$

If non-linearity is included in the Klein – Gordon equations or external fields which can effectively lead to the non-positive square of mass m^2, the separation of solutions having positive and negative frequencies into the two non-interacting theories disappears, and their unification is achieved in the framework of second quantization. Critical for its success is the causal spacetime structure which allows for the creation and propagation of physically observable particles only in one direction inside the light cone whereas negative-frequency modes travelling 'backward' in time are used to describe effects of their annihilation.

3.6 An alternative: third quantization or York gauge formalism?

Einstein's theory of gravity is essentially different from the relativistic particle theory, because the evolution in superspace $q^i = \big(g_{ab}(\mathbf{x}), \varphi(\mathbf{x})\big)$, an infinitely-dimensional analog of space x^α, is possible in all directions including those outside the 'light' cone of the DeWitt supermetric (due to indefiniteness of the potential term in the Wheeler – DeWitt equation). Therefore, bouncing off a caustic or at a turning point in superspace accompanied by a change in the direction of motion along the coordinate chosen as 'time' should not necessarily be interpreted as an exotic phenomenon like the creation or annihilation of the Universe. Simply, the choice of gauge which singles out this time coordinate in superspace was wrong: it is unable to globally parametrize the motion on both trajectory branches, before and after the bounce. In a class of gauges depending only on superspace coordinates (3.22), the solution of this problem is impossible because, as was shown before, caustics cause degeneracy of the Faddeev – Popov operator and lead to the problem of Gribov's copies. However, there is a gauge involving momenta which solves this problem.

This gauge was first introduced by York [72] and describes slicing of spacetime by a family of hyperspaces with the mean extrinsic curvature constant on each of them. The constant value of this mean curvature may be identified with time t, and as the trace of the extrinsic curvature coincides with the trace of gravitational momentum up to the scalar coefficient of weight unity, such a gauge has the form

$$\chi^\perp(q, p, t) = \frac{2}{3}\, g^{-1/2} g_{ab}\, p^{ab} - t = 0\,. \tag{3.89}$$

Note that this gauge is a scalar with respect to spatial diffeomorphisms generated by supermomentum constraints; therefore,

$$J^\perp_a = \{\chi^\perp, H_a\} \sim \chi^\perp\,, \tag{3.90}$$

and the Faddeev – Popov determinant factorizes on equations of gauge conditions

$$J\big|_{\chi^\perp = 0} = \det J^\perp_{\perp'}\, \det J^a_b\,. \tag{3.91}$$

Since $J^a_b = \{\chi^a, H_b\}$ does not introduce Gribov's copies related to the problem of time, one may forget gauges $\chi^a(q)$ fixing spatial diffeomorphisms and concentrate on gauge (3.89).

It is remarkable by the fact that the time variable it introduces on *phase* space,

$$T = \frac{2}{3}\, g^{-1/2} g_{ab}\, p^{ab}\,, \tag{3.92}$$

is likely to carry out unambiguous ADM reduction to the physical sector and globally cover both classical and quantum histories. This property was proved by York [72, 73] in a special parametrization of phase space variables which make use of the conformal properties of the metric. For sake of simplicity, we shall restrict ourselves to a pure gravitational theory without matter sources even though this formalism is easy to generalize to the case with matter. Let us consider conformal expansion of a 3-dimensional metric and its conjugate momenta

$$g_{ab} = \phi^4 \widetilde{g}_{ab}\,, \tag{3.93}$$

$$p^{ab} = \phi^{-4}\widetilde{p}^{\,ab} + \frac{1}{2}\widetilde{g}^{\,1/2}\widetilde{g}^{\,ab}\phi^2 T\,, \qquad \widetilde{g}_{ab}\widetilde{p}^{\,ab} = 0\,, \tag{3.94}$$

where $\widetilde{g}_{ab}$ is a conformally invariant metric built in a certain conformal gauge (a function of five independent variables), ϕ is the conformal factor, and $\widetilde{p}^{\,ab}$ is the traceless part of momentum. Transformation of the symplectic form

$$p^{ab}\dot{g}_{ab} = \widetilde{p}^{\,ab}\dot{\widetilde{g}}_{ab} - \widetilde{g}^{\,1/2}\phi^6\dot{T} + \frac{\mathrm{d}}{\mathrm{d}t}(\ldots) \tag{3.95}$$

shows that $\widetilde{p}^{\,ab}$ are conjugate to the components of the conformal metric $\widetilde{g}_{ab}$ while the momentum conjugate to T is equal to

$$P_T = -\widetilde{g}^{\,1/2}\phi^6 = -g^{1/2}\,. \tag{3.96}$$

The substitution of (3.93), (3.94) into the superhamiltonian constraint leads to the Lichnerowicz equation for the conformal factor

$$\widetilde{g}^{\,1/2}\left(\widetilde{\Delta} - \frac{1}{8}\widetilde{R}\right)\phi + \frac{1}{8}\frac{\widetilde{p}^{2}}{\widetilde{g}^{\,1/2}}\frac{1}{\phi^7} - \frac{3}{64}\widetilde{g}^{\,1/2}T^2\phi^5 = 0\,, \tag{3.97}$$

$$\widetilde{\Delta} = \widetilde{g}^{\,ab}\widetilde{\nabla}_a\widetilde{\nabla}_b\,, \qquad \widetilde{p}^{\,2} \equiv \widetilde{g}_{ac}\widetilde{g}_{bd}\widetilde{p}^{\,ab}\widetilde{p}^{\,cd} \tag{3.98}$$

in the conformal treatment of the initial value problem; as shown in Refs [72, 73], this equation 'almost everywhere' has the only limited positive solution, that is excepting the set of measure zero of physically unacceptable values of variables† $\widetilde{g}_{ab}$, $\widetilde{p}^{\,ab}$. The unique solution means that the linear operator obtained by varying (3.97) with respect to ϕ is non-degenerate; hence, the diagonal block of the Faddeev–Popov operator‡

$$J^{\perp}_{\perp'}\Big|_{H_\perp=0} = 2\phi^3\left\{\widetilde{\Delta} - \frac{1}{8}\widetilde{R} - \frac{7}{8}\frac{\widetilde{p}^{2}}{\widetilde{g}}\frac{1}{\phi^8} - \frac{15}{64}T^2\phi^4\right\}\frac{\delta(\mathbf{x},\mathbf{x}')}{\phi^3}\,,$$
$$\perp = \mathbf{x}\,, \qquad \perp' = \mathbf{x}' \tag{3.99}$$

is also non-degenerate which is in fact the necessary criterion for unambiguous ADM reduction. It follows from (3.95), (3.96) that the physical Hamiltonian in conformal phase space variables $\widetilde{g}_{ab}$, $\widetilde{p}^{\,ab}$ is obtained by substituting the solution of the Lichnerowicz equation $\phi\big[\widetilde{g}_{ab}(\mathbf{x}), \widetilde{p}^{\,ab}(\mathbf{x}), T(\mathbf{x})\big]$ into (3.96) and is numerically coincident with the three-dimensional volume of the Universe

$$H_{\text{phys}}(t)\big[\widetilde{g}_{ab}, \widetilde{p}^{\,ab}\big] = \int \mathrm{d}^3x\,\widetilde{g}^{\,1/2}\big(\phi[\widetilde{g}_{ab}, \widetilde{p}^{\,ab}, t]\big)^6 = \int \mathrm{d}^3x\, g^{1/2}. \tag{3.100}$$

The remaining reduction to physical variables in this Hamiltonian is achieved by imposing coordinate gauges on five independent components $\widetilde{g}_{ab}$ and their solution together with supermomentum constraints that acquire, in new variables (3.94), the form

$$\widetilde{\nabla}_a\widetilde{p}^{\,ab} = 0\,, \tag{3.101}$$

which eventually reduces the system to two independent degrees of freedom.

The fact that the global time (3.92) is defined by the conformal part of gravitational momenta suggests that Dirac quantization may also be carried out in new phase variables $(T, P_T;\, \widetilde{g}_{ab}, \widetilde{p}^{\,ab})$, with the 'conformal' superspace of variables $(T,\, \widetilde{g}_{ab})$ being chosen to serve as coordinates. Then, the constraint equation is the Lichnerowicz equation in which the conformal factor $\phi = (-P_T/\widetilde{g}^{\,1/2})^{1/6}$ plays the role of the momentum conjugate to T. Since it is non-polynomial in ϕ, this equation is very inconvenient from the viewpoint of the operator realization (recall that the formal operator realization which closes the algebra of quantum constraints exactly can be found when they are quadratic in momenta; see Section 3.2 and Refs [8, 66]). However, it turns out that by multiplying the Lichnerowicz equation by ϕ^7 and performing additional canonical transformation to variables $(\boldsymbol{\Phi}, \boldsymbol{\Pi};\, \widetilde{g}_{ab},\, \widetilde{\pi}^{ab})$

$$T = -\frac{2}{3}\widetilde{g}^{\,1/4}\frac{\boldsymbol{\Phi}}{\boldsymbol{\Pi}^{1/2}}\,, \qquad P_T = -\widetilde{g}^{\,-1/4}\boldsymbol{\Pi}^{3/2}\,, \tag{3.102}$$

$$\widetilde{p}^{\,ab} = \widetilde{\pi}^{ab} - \frac{1}{6}\boldsymbol{\Phi}\boldsymbol{\Pi}\widetilde{g}^{\,ab}\,, \tag{3.103}$$

it is once again possible to make the superhamiltonian constraint quadratic in all the momenta $(\boldsymbol{\Pi},\, \widetilde{\pi}^{ab})$

$$\mathcal{H}_\perp \equiv \phi^4 H_\perp = \frac{1}{\widetilde{g}^{1/2}}\bigg[2\boldsymbol{\Pi}\widetilde{\Delta}\boldsymbol{\Pi} - \frac{3}{2}(\widetilde{\nabla}\boldsymbol{\Pi})^2 - \widetilde{R}\boldsymbol{\Pi}^2 - \frac{1}{4}\boldsymbol{\Phi}^2\boldsymbol{\Pi}^2\bigg] + \frac{\widetilde{\pi}^{ab}\widetilde{\pi}_{ab}}{\widetilde{g}^{1/2}}\,. \tag{3.104}$$

Evidently, this expression is not only quadratic but also homogeneous in momenta, so that it assumes the form of the 'massless' equation§

$$\mathcal{H}_\perp(q,p) = \mathcal{G}^{ik}_\perp p_i p_k\,, \tag{3.105}$$

$$q^i = \big(\boldsymbol{\Phi}(\mathbf{x}),\, \widetilde{g}_{ab}(\mathbf{x})\big)\,, \qquad p_i = \big(\boldsymbol{\Pi}(\mathbf{x}),\, \widetilde{\pi}^{ab}(\mathbf{x})\big) \tag{3.106}$$

in condensed notations for conformal superspace, even though it contains the nonultralocal $\boldsymbol{\Phi\Phi}$-sector of the three-point function

$$\mathcal{G}^{ik}_\perp = \mathrm{diag}\,\big[\boldsymbol{G}(\mathbf{x}_\perp|\,\mathbf{x}_i, \mathbf{x}_k),\, g_{a(c}g_{bd)}\delta(\mathbf{x}_\perp, \mathbf{x}_i)\delta(\mathbf{x}_\perp, \mathbf{x}_k)\big]\,, \tag{3.107}$$

$$\boldsymbol{G}(\mathbf{x}_\perp|\,\mathbf{x}, \mathbf{x}') = \frac{1}{\widetilde{g}^{1/2}}\bigg[\delta(\mathbf{x}_\perp, \mathbf{x})\widetilde{\Delta}\delta(\mathbf{x}_\perp, \mathbf{x}') - \frac{3}{2}\widetilde{g}^{\,ab}\widetilde{\nabla}_a\delta(\mathbf{x}_\perp, \mathbf{x})\widetilde{\nabla}_b\delta(\mathbf{x}_\perp, \mathbf{x}') - \left(\widetilde{R} + \frac{1}{4}\boldsymbol{\Phi}^2\right)\delta(\mathbf{x}_\perp, \mathbf{x})\,\delta(\mathbf{x}_\perp, \mathbf{x}')\bigg]. \tag{3.108}$$

† In the full quantum theory, this set apparently falls into the integration range of quantum field values and therefore requires special consideration.

‡ Note that all the terms in this operator are negative-definite with the exception of the 3-dimensional scalar curvature in the conformal Laplacian $\widetilde{\Delta} - \widetilde{R}/8$; therefore, the proof of its non-degeneracy, trivial for $\widetilde{R} > 0$, requires in the opposite case more subtle conformal superspace methods [72, 73].

§ Note that this homogeneity property is analogous to that of the Hamiltonian constraint in the Ashtekar variables [74], although it is achieved without using the complex triad formulation and without increasing the number of constraints. Generally speaking, homogeneity in momenta disappears in the presence of matter. In the case of a conformally-invariant electromagnetic field, the superhamiltonian acquires a term linear in $\boldsymbol{\Pi}$.

Operator realization of the Hamiltonian constraint (3.105) in the conformal superspace of coordinates $(\boldsymbol{\Phi}(\mathbf{x}),\, \tilde{g}_{ab}(\mathbf{x}))$ is beyond the scope of the present paper. It is worthwhile to note only that in the main, it reduces to the replacement of the momentum $\boldsymbol{\Pi}(\mathbf{x})$ by the functional derivative $\boldsymbol{\Pi}(\mathbf{x}) = \hbar\delta/\mathrm{i}\delta\boldsymbol{\Phi}(\mathbf{x})$. The $\boldsymbol{\Pi}^2$ term in (3.105) being negative-definite, the resultant equation will again be of the hyperbolic type and will seemingly have solutions of different frequencies. However, semiclassical solutions having negative frequency ($\boldsymbol{\Pi} < 0$) are irrelevant because the phase-space momentum region (conformal factor)

$$\boldsymbol{\Pi} = \phi^4 \tilde{g}^{1/2} \tag{3.109}$$

belongs to the positive semiaxis. This explains the mechanism of the semiclassical selection of solutions having positive frequency for the Wheeler – DeWitt equation in conformal superspace which may also be realized at the non-perturbative level.

Thus, the York gauge formalism and the related conformal superspace resolves the problem of time as the problem of gravitational Gribov's copies. Third quantization seems to be unnecessary, however this concept and physics of Euclidean – Lorentzian transitions prove to be possible via a different mechanism. Indeed, the description of a classically forbidden state with the help of imaginary time implies complexification of the conformal superspace

$$T = -\mathrm{i}\mathcal{T} \tag{3.110}$$

(momentum and its trace become imaginary when the Euclidean space is sliced by hypersurfaces of the constant Euclidean time $\mathcal{T}(\mathbf{x}) = \tau$). This complexification makes the 'conformal' time variable

$$\boldsymbol{\Phi} = -\frac{3}{2}\,\phi^2 T = \frac{3}{2}\,\mathrm{i}\phi^2\mathcal{T}$$

purely imaginary and explicitly converts the Wheeler – DeWitt equation from hyperbolic to the elliptic one. This is the absolute analog of the transition from the hyperbolic to elliptic Klein – Gordon equation under the Wick rotation. In a conventional metric superspace, this analogy was implicit; specifically, transition to the Euclidean space did not imply a change of the DeWitt supermetric signature. On the contrary, in a conformal superspace, transition into the Euclidean region of spacetime is closely associated with its Euclidization achieved by the Wick rotation of the conformal mode

$$\boldsymbol{\Phi} \to -\mathrm{i}\boldsymbol{\Phi}\,. \tag{3.111}$$

Quantization in a conformal superspace requires a further study. Clearly, this formalism does not remove the possibility of third quantization and underbarrier phenomena with a change of the spacetime signature. But here, these phenomena are of different nature and have a somewhat different interpretation. In the first place, these differences are related to the behaviour in the vicinity of possible caustics in superspace. In a conformal superspace, caustics (if any) do not raise the problem of time and nor do they lead to a naive interpretation of one and the same Universe, considered prior to and after bouncing off a caustic, as the quantum superposition of two simultaneously existing states; in other words, the surface of constant (external) time (3.92) does not cross quantum histories more than once and does not give rise to Gribov's copies.

3.7 ADM quantization of the simplest minisuperspace model

The relativistic particle example considered in Section 3.5 is too simple and lacking many important features of real cosmological models (including the possibility of a dynamical change of the signature).

Let us illustrate the ADM reduction to physical variables for a two-dimensional minisuperspace (a, φ), where $a(t)$ is the scale factor of a closed homogeneous Universe with metric

$$\mathrm{d}s^2 = -N^2\,\mathrm{d}t^2 + a^2\,\mathrm{d}l^2\,, \tag{3.112}$$

$\varphi(t)$ is the spatially homogeneous mode of a scalar field with the minimal coupling and the self-interaction potential $U(\varphi)$. This theory is the standard field for modelling Universe creation (see Section 4); its minisuperspace action is

$$S = m_{\mathrm{P}}^2 \int a^3 \left[-\frac{\dot{a}^2}{a^2 N^2} + \frac{1}{m_{\mathrm{P}}^2}\frac{\dot{\varphi}^2}{2N^2} - \left(H^2(\varphi) - \frac{1}{a^2} \right) \right] \mathrm{d}t$$
$$= \int \mathrm{d}t\,\left[p_a\,\dot{a} + p_\varphi\dot{\varphi} - N\mathcal{H} \right], \tag{3.113}$$

where m_{P} is the Planck mass. The superhamiltonian constraint is

$$\mathcal{H} = -\frac{p_a^2}{4m_{\mathrm{P}}^2 a} + \frac{p_\varphi^2}{2a^3} + m_{\mathrm{P}}^2 a^3 \left(H^2(\varphi) - \frac{1}{a^2} \right) = 0\,. \tag{3.114}$$

Let us now perform the ADM reduction to physical variables in two different gauges (3.22).

(1) Let us first choose as the 'clock' the scale factor of the Universe or its function

$$\chi = a - f(t) = 0\,. \tag{3.115}$$

In accordance with the general formulas in Section 3.2, the Faddeev – Popov determinant $J_\perp^\perp \equiv J$, the lapse function $N^\perp \equiv N$, and the physical Hamiltonian (3.27) are expressed through the momentum p_a

$$J = \pm\frac{1}{2m_{\mathrm{P}}^2 a}|p_a|\,, \tag{3.116}$$

$$N = \frac{\mathrm{d}f}{\mathrm{d}t}\frac{1}{J} = \pm\frac{2m_{\mathrm{P}}^2 a}{|p_a|}\frac{\mathrm{d}f}{\mathrm{d}t}\,, \tag{3.117}$$

$$H_{\mathrm{phys}} = \pm\frac{\mathrm{d}f}{\mathrm{d}t}|p_a|\,, \tag{3.118}$$

where p_a is the function of physical variables obtained by the solution of the constraint equation (3.114) taking into account (3.115):

$$|p_a| = \frac{\sqrt{2}m_{\mathrm{P}}}{f(t)}\sqrt{m^2(\varphi, t) + p_\varphi^2}\,, \tag{3.119}$$

$$m^2(\varphi, t) \equiv m_{\mathrm{P}}^2 f^4 \left(H^2(\varphi) f^2 - 1 \right). \tag{3.120}$$

Quantum dynamics is defined by the action (3.26) [in the given case, coordinates of the physical phase space ξ and momenta conjugate to them π correspond to the pair $(\varphi,\, p_\varphi)$]:

$$S(\varphi,\, p_\varphi) = \int \mathrm{d}t \left[p_\varphi\dot{\varphi} - H_{\mathrm{phys}} \right] \tag{3.121}$$

and the Shrödinger equation (3.29):

$$i\hbar \frac{\partial}{\partial t}|\Psi\rangle = \pm \frac{df}{dt} \frac{\sqrt{2}m_P}{f(t)} \sqrt{m^2(\varphi, t) + p_\varphi^2}\, |\Psi\rangle \tag{3.122}$$

for the wave function $|\Psi\rangle$ with the conserved norm (3.30). We do not discuss here the problem of ordering the operators φ, p_φ in (3.119); in the generally accepted weak scalar field dependence of $H(\varphi)$, the corresponding correction terms are of the lowest order.

Here, the fundamental difference from the relativistic particle (Section 3.5) consists in the explicit time-dependence of parameters of the physical Hamiltonian, e.g. particle 'mass' (3.120). At $H^2a^2 < 1$, this theory describes a particle with imaginary mass, a tachyon. It is known that the solutions of classical dynamical equations obtained using the variation of the action (3.113) reveal [following the fixation of the reference frame (3.115)] strict correlation between φ and p_φ values at each moment. Quantum gravity must answer in the affirmative the following questions of special interest from the viewpoint of conformity with observations: How can the coherent state $\Psi(\varphi, t)$ with small dispersions and corresponding ($\varphi \leftrightarrow p_\varphi$) correlation be prepared? What is the spreading rate of this wave packet? The answer to the latter question is in the system dynamics. Specifically, the packet should be expected to spread extensively at the moment when 'mass' (3.120) vanishes even if the mean value of the momentum $p_\varphi \neq 0$ and the region $m^2 < 0$ is classically accessible (the expression under the root sign in (3.119) remains positive).

The Shrödinger equation (3.122) is usually considered in the semiclassical approximation. The quantum mechanical problem (3.122) evidently deserves a more detailed examination. In any case, it has the exact solution at $dm/d\varphi = 0$, i.e. when the moment p_φ is conserved. The fact that the sign of m^2 is indefinite in this problem is physically non-trivial.

The lapse function (3.117) and, consequently, proper time

$$\tau = \int_{t_1}^{t_2} N\, dt = \int_{a_1}^{a_2} \frac{2m_P^2 a}{|p_a|}\, da \tag{3.123}$$

between the states of the Universe with scale factors a_1, a_2 [invariant τ is independent of the choice of $f(t)$] are operators. In calculating the integral over time in (3.123), φ, p_φ should be regarded as time-dependent Heisenberg operators. In a semiclassical regime, when $m^2 \gg p_\varphi^2$, quantum dispersion of proper time is small, but it is certain to grow dramatically near the 'turning point'. If a_1, a_2 are localized at different sides of the turning point, with a_1 lying in the Euclidean region and a_2 in the Lorentzian one, then τ consists of two terms, imaginary and real, as was expected:

$$\tau = i\tau_E + \tau_L\,. \tag{3.124}$$

Note that the choice as the time variable of the scale factor itself [$a = t$, i.e. $f(t) = t$ in (3.115)] leads, in accordance with (3.117), to the expression for the g_{00}-component of the metric

$$g_{00} = -N^2 = -J^{-2} = -\frac{2m_P^2 a^4}{m^2 + p_\varphi^2}\,, \tag{3.125}$$

which has, in neglect of p_φ^2, a simple pole

$$g_{00} \sim \frac{1}{a - H^{-1}(\varphi)} \tag{3.126}$$

near the classical turning point. It is this expression [see Introduction, (1.7)] that Sakharov proposed to illustrate a change of the time signature.

(2) Let us now take as a 'clock' the scalar field φ, that is assume that

$$\chi = \varphi - t = 0\,. \tag{3.127}$$

Using the constraint equation (3.114) and following general formulas (3.19), (3.21), and (3.27), we have

$$J = \pm \frac{|p_\varphi|}{a^3}\,, \tag{3.128}$$

$$N = J^{-1} = \pm \frac{a^3}{|p_\varphi|}\,, \tag{3.129}$$

$$H_{\text{phys}} = \pm |p_\varphi|\,, \tag{3.130}$$

where

$$|p_\varphi| = \frac{a}{\sqrt{2}m_P} \sqrt{p_a^2 - U(a, t)}\,, \tag{3.131}$$

$$U(a, t) \equiv 4m_P^4 a^2 \left(H^2(t)a^2 - 1\right). \tag{3.132}$$

In this case, unlike the case of gauge (3.115), dependence of the physical Hamiltonian (3.130) on the 'coordinate' a is more important than its 'time'-dependence. At $U > 0$, we have an unstable situation reminiscent of an 'overturned' oscillator, i.e. roll-down of a semiclassical packet the centre of which moves along the classical trajectory $a(t)$ that corresponds to the evolution of the Lorentzian Universe in time. It is however incorrect to regard superspace regions with $U > 0$ as Lorentzian and those with $U < 0$ as Euclidean ones. In fact, a change of the signature (a change of sign N^2) occurs upon a change in the sign of the expression under the root in (3.131) which is not directly related to a change of sign of $U(a, t)$.

Here, like in the York gauge case, the use of a 'monotonically' changing variable as a clock eliminates the problem of Gribov's copies and requires a different interpretation of the signature change (see Section 3.6). The quantum-mechanical problem

$$i\hbar \frac{\partial \Psi(a, t)}{\partial t} = \frac{a}{\sqrt{2}m_P} \sqrt{-\hbar^2 \frac{\partial^2}{\partial a^2} - U(a, t)}\; \Psi(a, t) \tag{3.133}$$

is non-trivial. The meaning of a signature change in the language of solutions for Eqn (3.133) should be addressed in further studies.

Section 4 discusses the origin of the early Universe in the theory (3.113) in gauge (3.115), under the assumption that H weakly depends on φ and that, in accordance with the Hartle and Hawking 'no-boundary' prescription, the initial (at $a = 0$) state $\Psi(\varphi)$ is a state with zero momentum p_φ, i.e. φ-independent.

4. Quantum origin of the early Universe

The theory of the quantum origin of the early Universe is currently a most productive field for the application of physics of transitions with the changing spacetime signature. In the early 1980s, the synthesis of ideas of cosmological inflation [75, 40] and quantum state [17, 18, 60], creating initial conditions for the inflationary scenario, gave a powerful impact to the development of quantum cosmology. It eventually resulted in the invention of the Euclidean quantum gravity that brought scientists to the concept of third quantization, physics of wormholes, and the cosmological constant fixing mechanism first suggested by Coleman [30]. This chapter is devoted to quantum cosmology of the early Universe and gives examples illustrating the general formalism and problems considered in previous sections. A key problem of this theory which was not discussed above concerns a search for a state $\boldsymbol{\Psi}(q)$ that might be useful for the description of an early, essentially quantum Universe whose evolution leads to the present large-scale classical picture of the world.†

A particular problem to be solved in the framework of quantum cosmology of the early Universe is the demonstration of the validity of the semiclassical expansion at the initial quantum stage of inflation and the calculation of its energy scale. The inflationary paradigm is especially attractive because it allows one to avoid unreliable predictions of quantum gravity, by making use of sub-Planckian physics with the characteristic value of the Hubble constant $H = \dot{a}/a \sim 10^{-5} m_{\mathrm{P}}$ much lower than the Planck mass $m_{\mathrm{P}} = G^{1/2}$. The predictions of the inflation theory strongly depend on this energy scale which must be chosen in such a way as to ensure a sufficiently large parameter of exponential expansion $\exp N \sim \exp(60)$ during the inflationary stage and generate the necessary level of perturbations to create a contemporary large-scale structure. However, this scale in the inflation theory is a free parameter and can hardly be fixed without invoking the ideas of quantum gravity and quantum cosmology. These ideas imply that the quantum state of the Universe in a semiclassical regime gives rise to an ensemble of inflationary universes with different values of the Hubble constant which later evolve approximately according to classical equations of motion. The quantum state allows one to find the distribution function of such an ensemble and to interpret its probable maximum (if any) as generating the energy scale of inflation. We shall demonstrate that although the realization of such an idea in the tree-level approximation of quantum cosmology has failed [77, 78], it proves to be possible in the one-loop approximation [79, 8, 80, 69] which requires to use in full the quantum gravity formalism discussed in previous sections.

4.1 Hartle – Hawking and Vilenkin quantum states as a source of inflationary Universe

Today, the inflationary stage in the early Universe is widely accepted to fairly well explain the problem of the creation of a contemporary large-scale structure of the observable part of spacetime and its microwave background [40, 75]. The inflationary stage is a period in dynamics of the early Universe described by the de Sitter or quasi-de Sitter geometry created by the effective cosmological constant Λ which is in its turn generated by other slowly varying fields. For example, in the chaotic inflation model with the scalar inflaton field ϕ minimally coupled to the metric tensor $G_{\mu\nu}$

$$L(G_{\mu\nu}, \phi) = G^{1/2} \left\{ \frac{m_{\mathrm{P}}^2}{16\pi} R(G_{\mu\nu}) - \frac{1}{2} (\nabla\phi)^2 - U(\phi) \right\}, \quad (4.1)$$

† The possibility of the creation of the initial inhomogeneity spectrum by primary quantum fluctuations has been discussed in one of the Sakharov's pioneering works [76].

in the approximation of the slow roll down the potential barrier $U(\phi)$ (assumed to be monotonically growing with ϕ) when the rate of time dependence of ϕ is much lower than the inflation rate of the scale factor a (Hubble constant $H = \dot{a}/a$), equations of motion assume the form

$$\dot{\phi} \simeq -\frac{1}{3H} \frac{\partial U}{\partial \phi} \ll H\phi \,, \quad (4.2)$$

$$H = H(\phi) \simeq \sqrt{\frac{8\pi U(\phi)}{3m_{\mathrm{P}}^2}} \,, \quad (4.3)$$

and the effective cosmological constant $\Lambda = 3H^2$ is defined by the inflaton field potential. It remains approximately constant throughout the inflationary stage due to the slow decrease of ϕ and essentially decreases only at the end of this stage. This leads to the decay of the effective cosmological constant into inflaton oscillations the energy of which leads to the reheating of matter in the Universe and its transition first to the radiation-dominated and then to the matter-dominated stage. The objective of quantum cosmology is to prepare, at the quantum level, the necessary initial data for such a picture by choosing an appropriate quantum state of the Universe $\boldsymbol{\Psi}(q)$ the general theory of which has been considered above. One of the fruitful ideas for the realization of this task which dates to the pioneering works of Hartle – Hawking and Vilenkin [17, 18, 60] consists in the fact that such initial data emerge as a result of quantum tunnelling described by the wave function $\boldsymbol{\Psi}(q)$ and representing a transition with changing spacetime signature.

In the context of closed cosmology, de Sitter Lorentzian spacetime may be considered as a result of quantum tunnelling from the classically forbidden domain described by the de Sitter Euclidean geometry. A simple picture of tunnelling geometry illustrating such a mechanism is shown in Fig 3. The de Sitter solution of Einstein's equations with the cosmological constant $\Lambda = 3H^2$

$$\mathrm{d}s_{\mathrm{L}}^2 = -\,\mathrm{d}t^2 + a_{\mathrm{L}}^2(t)\, c_{ab}\, \mathrm{d}x^a\, \mathrm{d}x^b \,, \quad (4.4)$$

$$a_{\mathrm{L}}(t) = \frac{1}{H} \cosh\,(Ht) \quad (4.5)$$

describes the expansion of a spherical hypersurface with the metric of a 3-dimensional sphere c_{ab} and the scale factor $a_{\mathrm{L}}(t)$.

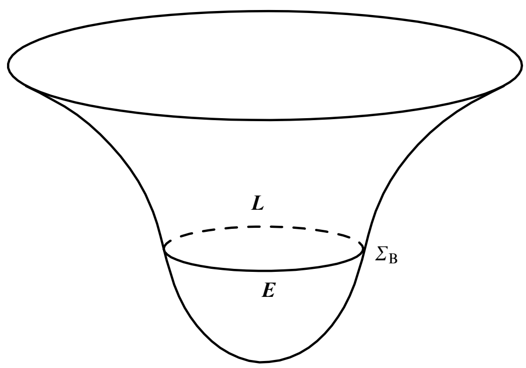

Figure 3. Graphical representation of the Lorentzian spacetime $\boldsymbol{L}$ nucleating from the Euclidean manifold $\boldsymbol{E}$, with the topology of a 4-dimensional ball, at the bounce surface Σ_{B}. This construction is used in the Hartle – Hawking prescription for the cosmological quantum state.

Its Euclidean analogue with the de Sitter metric of the positive signature

$$ds^2 = d\tau^2 + a^2(\tau)\, c_{ab}\, dx^a\, dx^b\,, \tag{4.6}$$

$$a(\tau) = \frac{1}{H}\sin(H\tau) \tag{4.7}$$

describes geometry of a 4-dimensional sphere with radius $R = 1/H$ and 3-dimensional sections parametrized by the latitude angle $\theta = H\tau$. The two metrics are related by the analytic continuation to the complex plane of the Euclidean time τ [81, 82]

$$\tau = \frac{\pi}{2H} + it\,, \qquad a_L(t) = a\left(\frac{\pi}{2H} + it\right). \tag{4.8}$$

This analytic continuation may be interpreted as a quantum tunnelling from the Lorentzian spacetime to the Euclidean one and is shown in Fig. 3 as two manifolds (4.4)–(4.7) matched together along the equatorial section of the 4-dimensional sphere $\tau = \pi/2H$ $(t = 0)$ — the bounce surface Σ_B.

The two known quantum states that semiclassically give rise to such a mechanism for the creation of inflationary universes are the Hartle–Hawking wave function [17, 18] and the Vilenkin tunnelling wave function [60]. In the approximation of two-dimensional minisuperspace consisting of two variables, the scale factor a and the inflaton scalar field ϕ,

$$q^i = (a, \phi)\,, \tag{4.9}$$

these wave functions $\boldsymbol{\Psi}_{NB}(a, \phi)$ and $\boldsymbol{\Psi}_T(a, \phi)$ satisfy the minisuperspace Wheeler–DeWitt equation and, in semiclassical terms, represent its linearly independent solutions

$$\boldsymbol{\Psi}_{NB}(a, \phi) \sim \exp\left[-\boldsymbol{I}(a, \phi)\right], \qquad \boldsymbol{\Psi}_T(a, \phi) \sim \exp\left[+\boldsymbol{I}(a, \phi)\right], \tag{4.10}$$

in which the Euclidean Hamilton–Jacobi function $\boldsymbol{I}(a, \phi)$ is evaluated on a particular family of solutions for classical Euclidean equations of motion satisfying special Hartle–Hawking boundary conditions at $a = 0$ and boundary conditions (a, ϕ) at the end point, the argument of this function. At $a = 0$, the scalar field derivative with respect to the Euclidean time τ must be zero and $da/d\tau = 1$ (τ measures the proper distance). This is equivalent to the requirement of regularity of the 4-metric in the vicinity of the pole of a 4-dimensional sphere (4.6), (4.7) at $\tau = 0$. In the leading order of the slow roll approximation, when the inflaton field is constant, such a solution just coincides with this metric at the Hubble constant (4.3) while the Hamilton–Jacobi function equals

$$\boldsymbol{I}(a, \phi) = -\frac{\pi m_P^2}{2H^2}\left\{1 - \left[1 - H^2(\phi)a^2\right]^{3/2}\right\},$$

$$H^2(\phi) = \frac{8\pi U(\phi)}{3m_P^2}\,. \tag{4.11}$$

Till the point (a, ϕ) remains in the two-dimensional superspace region under the curve (see Fig. 4)

$$a = \frac{1}{H(\phi)}\,, \tag{4.12}$$

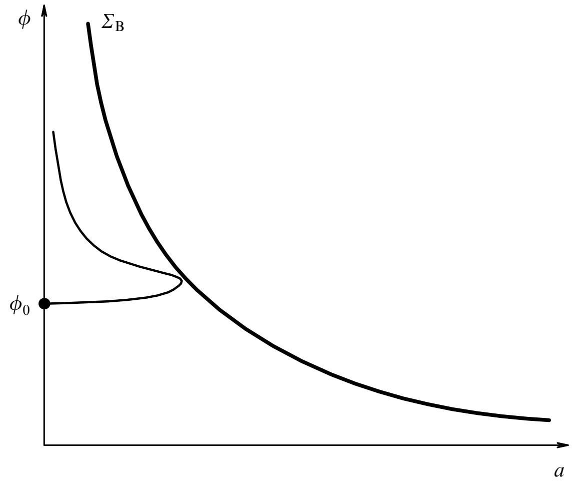

Figure 4. Two-dimensional minisuperspace of the scale factor a and the inflaton scalar field ϕ in the chaotic inflation model. Euclidean extremals (in the slow roll approximation) start at $a = 0$ and large initial values of ϕ_0, with the Hartle–Hawking boundary condition $d\phi/da|_{a=0} = 0$ in the form of trajectories reflected from caustic Σ_B, $a \simeq 1/H(\phi)$ and running into the region $\phi \to \infty$, $a \to 0$. Their analytic continuation into the plane of complex time $\tau \simeq \pi/2H + it$ gives rise to classical trajectories which describe Lorentzian spacetime. The caustic Σ_B in superspace is a boundary of transition with the changing spacetime signature and is formed by a one-dimensional family of zeroes of the Faddeev–Popov 'determinant' $J = 0$ although it does not contain zeroes of Eqn (3.73) in gauge (4.22), in perfect agreement with the discussion in Section 3.

the Universe is in the underbarrier state described by the Euclidean spacetime with the metric of a 4-dimensional sphere. Euclidean extremals originating at $a = 0$ have caustic† (4.12) and can not enter region $a > 1/H(\phi)$. However, they can be extended to this region in the complex time (4.8), and the Euclidean function $\boldsymbol{I}(a, \phi)$ acquires the Lorentzian part

$$\boldsymbol{I}(a, \phi) = \boldsymbol{I}(\phi) \pm i\boldsymbol{S}(a, \phi)\,, \qquad a > \frac{1}{H(\phi)}\,, \tag{4.13}$$

$$\boldsymbol{S}(a, \phi) = -\frac{\pi m_P^2}{2H^2}\left[H^2(\phi)\, a^2 - 1\right]^{3/2}. \tag{4.14}$$

Here, $\boldsymbol{I}(\phi)$ is the Euclidean action of a theory with the Lagrangian (4.1) calculated on a gravitational semi-instanton, i.e. hemisphere (4.6) $(0 \leqslant \tau \leqslant \pi/2H)$

$$\boldsymbol{I}(\phi) = -\frac{3m_P^4}{16U(\phi)}\,. \tag{4.15}$$

This action defines the amplitude of the wave functions (4.10) in the classically-allowed region

$$\boldsymbol{\Psi}_{NB}(a, \phi) \sim \exp\left[-\boldsymbol{I}(\phi)\right]\cos\left[\boldsymbol{S}(a, \phi) + \frac{\pi}{4}\right], \tag{4.16}$$

$$\boldsymbol{\Psi}_T(a, \phi) \sim \exp\left[+\boldsymbol{I}(\phi) + i\boldsymbol{S}(a, \phi)\right], \qquad a > \frac{1}{H(\phi)}\,, \tag{4.17}$$

† In the lowest order of the slow roll approximation, the problem is actually a one-dimensional one. Therefore, Eqn (4.12) represents just a set of turning points, although beyond this approximation the curve really turns out to be the envelope of a family of Euclidean trajectories [83].

which is interpreted in the tree-level (the lowest in $\hbar$) approximation as a distribution function of the ensemble of inflationary universes described by the Hamilton–Jacobi function (4.14) in Lorentzian spacetime. The main parameter characterizing these universes is the value of the effective Hubble constant $H = H(\phi)$ or the scalar curvature of de Sitter space the distribution over which is contained in the corresponding Hartle-Hawking $\rho_{\rm NB}(\phi)$ [17] and Vilenkin $\rho_{\rm T}(\phi)$ [84] functions:

$$\rho_{\rm NB}(\phi) \sim \exp\left[-2\boldsymbol{I}(\phi)\right], \qquad \rho_{\rm T}(\phi) \sim \exp\left[+2\boldsymbol{I}(\phi)\right]. \tag{4.18}$$

The Hartle–Hawking wave function differs from the Vilenkin wave function (as well as their distribution functions do) by the boundary conditions in superspace: while the tunnelling state $\boldsymbol{\Psi}_{\rm T}(a, \phi)$ at $a > 1/H(\phi)$ contains only an outgoing wave and describes an expanding Universe, the Hartle–Hawking wave function $\boldsymbol{\Psi}_{\rm NB}(a, \phi)$ in the Lorentzian region is a superposition of oppositely evolving cosmologies which may be interpreted, in the context of the discussion in Sections 2 and 3, as components with different frequencies of the solution of the Wheeler–DeWitt equation. The tunnelling wave function is defined with the help of the aforementioned boundary conditions of an outgoing wave in the Lorentzian region of superspace and an additional condition of ϕ-independence of $\boldsymbol{\Psi}_{\rm T}(a, \phi)$ at $a = 0$ [78, 85]. There is a more fundamental and model-independent prescription for the Hartle–Hawking wave function in the form of a functional integral over Euclidean geometries [17, 18] which, in the tree-level approximation, leads to expression (4.16) as a dominant contribution of the saddle point of this integral, i.e. the Euclidean–Lorentzian extremal (4.4)–(4.7).

The distribution functions $\rho_{\rm NB}(\phi)$ and $\rho_{\rm T}(\phi)$ describe opposite results of the most probable underbarrier tunnelling to the minimum and maximum of the inflaton potential $U(\phi) \geqslant 0$, respectively (although the minimum $U(\phi) = 0$ is, generally speaking, beyond the scope of applicability of the slow roll approximation). The above equations hold for the model (4.1), but they equally apply to the theory with non-minimal coupling of the scalar inflaton φ

$$\boldsymbol{L}(g_{\mu\nu}, \varphi) = g^{1/2}\left\{\frac{m_{\rm P}^2}{16\pi}\, R(g_{\mu\nu}) - \frac{1}{2}\,\xi\varphi^2 R(g_{\mu\nu}) - \frac{1}{2}(\nabla\varphi)^2 - \frac{1}{2}\, m^2\varphi^2 - \frac{\lambda}{4}\,\varphi^4\right\}, \tag{4.19}$$

provided that $L(G_{\mu\nu}, \phi)$ is considered to be Einstein's parametrization of Lagrangian $\boldsymbol{L}(g_{\mu\nu}, \varphi)$, with fields

$$(G_{\mu\nu}, \phi) = \left(\left(1 + \frac{8\pi|\xi|\varphi^2}{m_{\rm P}^2}\right) g_{\mu\nu}, \phi(\varphi)\right),$$

being related to $(g_{\mu\nu}, \varphi)$ through known conformal transformations [86–88]. Given a negative constant of the non-minimal coupling $\xi = -|\xi|$, this model easily leads to the chaotic inflationary scenario [89] with the inflaton potential in the Einstein parametrization

$$U(\phi)\Big|_{\phi=\phi(\varphi)} = \frac{m^2\varphi^2/2 + \lambda\varphi^4/4}{\left(1 + 8\pi|\xi|\varphi^2/m_{\rm P}^2\right)^2}, \tag{4.20}$$

including the case of a spontaneously broken symmetry at scale v ($m^2 = -\lambda v^2 < 0$) in the Higgs potential $\lambda(\varphi^2 - v^2)^2/4$. At greater φ, the potential (4.20) becomes constant and displays two types of behaviour at intermediate values of the inflaton field, depending on the parameter

$$\delta \equiv -\frac{8\pi|\xi|m^2}{\lambda m_{\rm P}^2} = \frac{8\pi|\xi|v^2}{m_{\rm P}^2}.$$

At $\delta > -1$, it has no local maxima and generates a slow decrease of the scalar field leading to a standard scenario with the finite duration of the inflationary stage, whereas for $\delta < -1$, it has the local maximum at

$$\bar{\varphi} = \frac{m}{\sqrt{\lambda|1 + \delta|}}$$

and, owing to the negative slope of the potential, leads to inflation of unlimited duration for all models with the scalar field growing from the initial value $\varphi_{\rm I} > \bar{\varphi}$.

The tree-level distribution functions (4.18) for such a potential do not suppress over-Planckian scales and cannot be normalized at large φ:

$$\int^{\infty} {\rm d}\varphi\, \rho_{\rm NB,\,T}(\varphi) = \infty,$$

which calls in question the validity of the semiclassical expansion. It is only for a tunnelling wave function at $\delta < -1$ that the distribution $\rho_{\rm T}(\phi)$ has a local peak at $\bar{\varphi}$ which may serve as a source of the most probable energy scale of inflation at reasonable sub-Planckian values of the Hubble constant. However, this peak requires large positive mass of the inflaton field $m^2 > \lambda m_{\rm P}^2/(8\pi|\xi|)$ which is too big for the reasonable values $\xi = -2 \times 10^4$, $\lambda = 0.05$ [86] and formally generates the infinite duration of the inflationary stage (since it starts from inflaton potential maximum).

4.2 One-loop distribution function of inflationary universes

It is worth noting that the above calculation of the tree-level distribution functions did not practically require to know the correct probabilistic inner product of cosmological wave functions which was given so much attention in Sections 2 and 3. It was sufficient to calculate and square the wave function amplitude which, due to specific features of the model, turned out to be a function on the section of a two-dimensional minisuperspace transversal to the coordinate a which normally plays the role of time. Therefore, the distribution function thus obtained proved to be defined on a physical subspace of correct dimensionality, i.e. one-dimensional space of the inflaton field. The situation changes beyond the tree-level approximation: calculations are impossible without knowing both the correct inner product and the wave function with the pre-exponential factor in the required approximation. Also one needs to go beyond the minisuperspace approximation because now the distribution function includes a non-trivial contribution of the integration over virtual quantum fields frozen in the tree-level approximation. In the one-loop order, which we restrict ourselves with, these fields can be taken into account in a linear approximation. As before, the main approximation in the theory of chaotic inflationary Universe consists in the minisuperspace model with the scale factor a and the spatially homogeneous scalar inflaton φ whereas inhomogeneous fields of all possible spins are considered as perturbations on this background. Taken together, they form the superspace of variables

$$q^i = \left(a,\, \varphi,\, \varphi(\mathbf{x}),\, \psi(\mathbf{x}),\, A_a(\mathbf{x}),\, \psi_a(\mathbf{x}),\, h_{ab}(\mathbf{x}), \ldots\right). \tag{4.21}$$

In order to calculate and interpret distribution functions, we shall need specific reduction to ADM variables ξ as described in Section 3. This reduction is reasonable to carry out separately in the minisuperspatial sector of full superspace (a, φ) and in the inhomogeneous mode sector. Let us choose, as above, an inflaton field φ as a physical variable whose distribution function needs to be calculated, taking the solution of classical equations of motion (4.5) with $H = H(\varphi)$ as a gauge

$$\chi^{\perp}(a, \varphi, t) = a - \frac{1}{H(\varphi)} \cosh\left[H(\varphi)t\right] = 0\,, \qquad (4.22)$$

which simultaneously plays the role of the parametrization of minisuperspace coordinates in terms of physical variable φ.† The ADM reduction for linearized inhomogeneous field modes reduces to the choice of their transversal (T) and transvers-traceless components (TT) which results in the following complete set of physical variables

$$\xi^A = (\varphi, f)\,, \quad f = \left(\varphi(\mathbf{x}), \psi(\mathbf{x}), A_a^{\mathrm{T}}(\mathbf{x}), \psi_a^{\mathrm{T}}(\mathbf{x}), h_{ab}^{\mathrm{TT}}(\mathbf{x}), \ldots\right). \qquad (4.23)$$

The ADM reduction at the quantum level is easy to perform as described in Section 3 for the tunnelling state (4.17). However, it encounters serious difficulty when applied to the Hartle–Hawking function (4.16) which contains both positive and negative frequencies and gives rise, in gauge (4.22), to Gribov's copies corresponding to components with different frequencies. It has been shown in Section 3.6 above that these copies are artifacts of using an inappropriate gauge condition the surface of which crosses the classical extremal (4.5) of one and the same Universe twice, before and after bouncing off the minimal value of the scale factor $a = 1/H(\varphi)$. This results in interpreting the components with different frequencies in (4.16) as superpositions of two *simultaneously* existing states of an expanding and contracting Universe. It is possible to overcome this problem by quantization in the conformal superspace of Section 3.6, whereas in the semiclassical approximation it is sufficient to consider the quantum ADM reduction for individual positive-frequency (or negative-frequency) components (4.16) and obtain the corresponding wave function of physical variables $\Psi(\xi, t) = \Psi(\varphi, f|t)$. Then, the distribution function of the inflaton field φ should be regarded as a diagonal element of the pure state density matrix $\mathrm{tr}_f|\Psi\rangle\langle\Psi|$. It can be obtained from $|\Psi\rangle = \Psi(\phi, f|\,t)$ by averaging over the remaining physical field modes f

$$\rho\,(\phi\,|\,t) = \int \mathrm{d}f\, \Psi^*(\phi, f|\,t)\Psi(\phi, f|\,t)\,, \qquad (4.24)$$

and this does not simply reduce to squaring the wave function.

The calculation of the Hartle–Hawking and Vilenkin wave functions, both in the one-loop order in $\hbar$ and perturbatively in inhomogeneous modes f (4.23) on the background of the Robertson–Walker metric, has been reported by many authors [90, 82, 91, 85, 79, 69, 80]. It may use either the functional integration over regular fields in Euclidean space with metric (4.6), (4.7) or the generic one-loop kernel (3.55) (more specifically, its Euclidean analog) by composing it with a special wave function at the initial moment.‡ The two wave functions are actually a Gaussian function of variables f — their Euclidean vacuum invariant with respect to the de Sitter group.§ This makes integration over f in (4.24) trivial and leads to the fundamental algorithm for the one-loop distribution function which is valid for both the Hartle–Hawking quantum state [79, 8, 69, 93, 80] and the tunnelling quantum state [92]

$$\rho_{\mathrm{NB,\,T}}(\varphi) \cong \frac{1}{H^2(\varphi)} \exp\left[\mp 2\boldsymbol{I}(\varphi) - \boldsymbol{\Gamma}_{\,1\text{-loop}}(\varphi)\right]. \qquad (4.25)$$

It turns out that the one-loop corrections to the both functions are identical and largely depend on the Euclidean effective action of the complete system of quantum fields $\xi(x)$

$$\boldsymbol{\Gamma}_{\,1\text{-loop}}(\varphi) = \frac{1}{2}\,\mathrm{tr}\left(\ln \frac{\delta^2 I[\xi]}{\delta\xi(x)\,\delta\xi(y)}\bigg|_{\mathrm{DS}}\right). \qquad (4.26)$$

The effective action is computed on the (quasi-)de Sitter gravitational instanton, a 4-dimensional sphere with radius $1/H(\varphi)$, and is therefore a function of φ. Such closed Euclidean manifold is obtained by doubling a semi-instanton [69], that is by matching together two hemispheres along the equatorial hypersurface Σ_{B} (at which quantum transition with the changing signature occurs). The procedure for calculating the distribution function is graphically depicted in Fig 5. The wave function and its conjugate involved in the

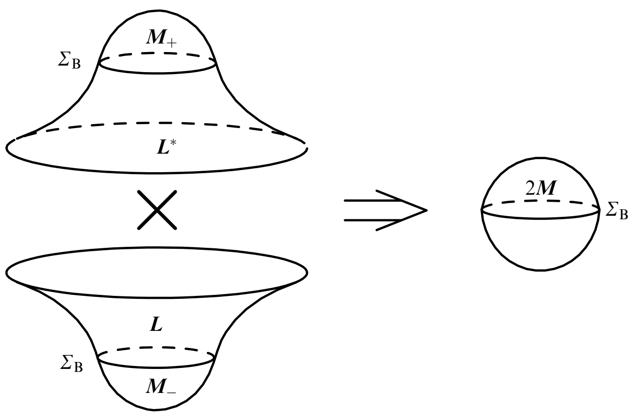

Figure 5. Graphical representation of the calculation of the quantum distribution function for Lorentzian universes. Composition of Euclidean-Lorentzian spacetime $\boldsymbol{M}_- \cup \boldsymbol{L}$ and its complex-conjugate and oppositely oriented copy $\boldsymbol{M}_+ \cup \boldsymbol{L}^*$ leads to the doubled Euclidean manifold $2\boldsymbol{M}$, a gravitational instanton, which carries the Euclidean effective action of the theory. Mutual cancellation of Lorentzian regions $\boldsymbol{L}$ and $\boldsymbol{L}^*$ reflects unitarity of the theory in physical spacetime with the Lorentzian signature.

† This gauge is very convenient because in the semiclassical approximation it corresponds to the choice of the proper time with the lapse function $N^{\perp} = 1$ [8].

‡ The Euclidean analog of (3.55) can be obtained by replacing the Lorentzian Hamilton–Jacobi function by the Euclidean one with Euclidean times τ and τ'. The Hartle–Hawking and Vilenkin wave functions are derived from this expression by shrinking the spatial section at $\tau' = 0$ to a point and integrating over all values of ξ' with a unit measure.

§ This is a Gaussian function also in the Euclidean region, but its dispersion obtained from basis functions of linearized Euclidean equations of motion [69, 92] tends to infinity at $\tau \to 0$. This corresponds to the infinite dispersion at the initial moment for the wave function involved in constructing the Hartle–Hawking and Vilenkin quantum states by means of the above composition with the kernel of the Euclidean evolution.

scalar product (4.24) can be represented by two Euclidean-Lorentzian manifolds; in calculating the inner product, the contributions of Lorentzian regions are mutually cancelled due to unitarity of the theory and the remnant is the Euclidean effective action computed on a closed instanton, obtained by gluing two hemispheres of the above type [69, 93].

Formula (4.25) possesses a number of remarkable properties. Note that, since we started from the quantum state of the ADM physical variables in a certain gauge and used it to calculate the distribution function (which seemingly had to be dependent on the choice of this gauge), the final result is gauge-invariant and gauge-independent. For the classical action (4.25), this holds by definition whereas the effective action $\boldsymbol{\Gamma}_{\text{1-loop}}(\varphi)$ does not depend on the choice of gauge condition because it is calculated on the mass shell [6, 94], and the pre-exponential factor $1/H^2(\varphi)$ is virtually gauge-invariant since it is expressed through the scalar curvature† $R(g_{\mu\nu}) = 4\Lambda = 12H^2$. Thus, the function (4.25) describes the distribution of a gauge-invariant observable, i.e. the scalar curvature in a quantum ensemble of de Sitter universes. This accounts for its gauge-invariance and gauge-independence.

Another property of the algorithm (4.25) is related to unitarity of the theory. Note that formula (4.25) is given for the distribution function of universes at the initial moment of their quantum nucleation from the Euclidean semi-instanton $t = 0$. The distribution function (4.24) is derived from (4.25) in accordance with the relation

$$\rho\,(\phi\,|\,t) = \rho_{\mathrm{NB,T}}\big(\varphi(\phi, t)\big)\left|\frac{\partial\varphi(\phi, t)}{\partial\phi}\right|, \tag{4.27}$$

where $\varphi(\phi, t)$ is the field value at this nucleation as a function of field ϕ to which the Universe evolves by the moment t [whose distribution function $\rho\,(\phi\,|\,t)$ is being calculated] [8, 69]. This relation proves unitarity of the theory, that is a conservation of the full probability $\int \mathrm{d}\phi\,\rho(\phi\,|\,t) = \text{const}$. The transformation Jacobian $\left|\partial\varphi(\phi, t)/\partial\phi\right| = 1/u_\varphi(t)$ is expressed through the linearized mode (basis function) of the inflaton field [see Eqn (3.71) in Section 3 in which ϕ and φ play the roles of ξ and ξ', respectively]. Both this Jacobian and the basis function remain regular on the caustic in two-dimensional minisuperspace separating the Euclidean and Lorentzian regions, but the corresponding Faddeev – Popov determinant degenerates, which serves to illustrate the general discussion in Section 3. To sum up, the full distribution function decomposes into the gauge-dependent, non-invariant Jacobian $\left|\partial\varphi(\phi, t)/\partial\phi\right|$ which ensures explicit unitarity of the theory and the gauge-invariant distribution function for gravitational instantons (4.25) which accumulates quantum corrections in a manifestly covariant form.

The latter property follows from the fact that the effective action (4.25) calculated in physical variables can be identically transformed to the manifestly covariant form (in the background-fields covariant gauge [6, 94]). Then, it can be calculated by covariant methods, e.g. by covariant regularization of its ultraviolet divergences. Only this guarantees correct calculation of the high-energy scaling behaviour $\boldsymbol{\Gamma}_{\text{1-loop}}(\varphi)$ and distribution functions. Indeed, in the high-energy limit of a large inflaton field which corresponds to the Hubble constant limit

$$H(\varphi) \simeq \sqrt{\frac{\lambda}{12|\xi|}}\,\varphi \to \infty$$

in the model (4.19), the effective action is calculated and renormalized on the de Sitter instanton of vanishing size H^{-1}. Therefore, it is asymptotically defined by the total anomalous scaling behaviour Z of the theory on such a space

$$\boldsymbol{\Gamma}_{\text{1-loop}}\Big|_{H\to\infty} \simeq Z \ln\frac{H}{\mu}\,. \tag{4.28}$$

Here μ is the renormalization mass parameter or the dimensional cut-off parameter generated by the finite fundamental string theory provided that the model (4.19) is considered as its effective sub-Planckian limit.

In the one-loop approximation, the parameter Z is defined by the second DeWitt coefficient [6, 94] of all covariant field multiplets integrated over the de Sitter instanton volume

$$Z = \frac{1}{16\pi^2}\int_{\mathrm{DS}} \mathrm{d}^4x\, g^{1/2} a_2(x, x)\,, \tag{4.29}$$

and is critically dependent on the phenomenological model of particles which includes the Lagrangian (4.19) as the graviton-inflaton sector. This quantity determines the complete set of one-loop logarithmic divergences and a set of corresponding β-functions.

The use of Eqns (4.25) and (4.28), with account for $H \sim \varphi(\varphi \to \infty)$, shows that the quantum distribution function, unlike its tree-level approximation, contains a Z-dependent factor

$$\rho_{\mathrm{NB,T}}(\varphi) \cong \frac{1}{\varphi^{Z+2}}\exp\big[\mp 2\boldsymbol{I}(\varphi)\big]\,. \tag{4.30}$$

This modification can make the Hartle – Hawking and Vilenkin wave functions normalizable at over-Planckian energies provided that parameter Z satisfies the inequality

$$Z > -1\,, \tag{4.31}$$

which serves as the criterion for the choice of a consistent field model of particles in the early Universe and is a conclusive argument in favour of the applicability of the semiclassical loop expansion [79, 95]. Although, strictly speaking, Eqn (4.30) holds only in the limit $\varphi \to \infty$, it may equally be used for the qualitative description at intermediate energies. Distribution of (4.30) in this region can give rise to the probability peak at $\varphi = \varphi_I$ with quantum dispersion σ, $\sigma^{-2} = -\,\mathrm{d}^2\ln\rho(\varphi_I)/\mathrm{d}\varphi_I^2$:

$$\varphi_I^2 = \frac{2|I_1|}{Z+2}\,,\qquad \sigma^2 = \frac{|I_1|}{(Z+2)^2}\,,$$

$$I_1 = -24\pi\,\frac{|\xi|}{\lambda}\,(1+\delta)m_{\mathrm{P}}^2\,, \tag{4.32}$$

† It is known that the theory of gravity has no local gauge invariants with respect to diffeomorphism transformations because even spatial scalars with respect to these diffeomorphisms undergo transformation by the term containing the Lie derivative. However, in a given setting of the problem, the scalar curvature on a homogeneous (quasi-)de Sitter space is approximately constant; hence, it may be regarded as a local invariant within the approximation used in the present paper.

where I_1 is the second coefficient of expansion of the Euclidean action in inverse powers of φ,

$$2\boldsymbol{I}(\varphi) = I_0 + \frac{I_1}{\varphi^2} + O\left(\frac{1}{\varphi^4}\right).$$

For the Hartle–Hawking quantum states and the tunnelling state, this peak is realized in complementary ranges of the parameter δ. In the case of the Hartle–Hawking state, it exists only at $\delta < -1$ $(I_1 > 0)$ and thus corresponds to the infinite duration of inflation with the field φ located at the negative slope of the inflaton potential (4.20) and infinitely growing from the initial value $\varphi_I > \bar{\varphi}$. For the tunnelling state, this peak occurs at $\delta > -1$ and leads to the finite duration of the inflationary stage with the parameter of the exponential expansion [80]

$$N(\varphi_I) = \frac{8\pi^2|\xi|\big(1+6|\xi|\big)}{\lambda(Z+2)}.$$

We shall examine this case in the following section because it describes the generally recognized inflationary scenario with the radiative-dominant stage following inflation.

4.3 Non-minimal inflation and particle physics in the early Universe

The status of the inflation theory was recently confirmed by observations of cosmic microwave background in satellite experiments COBE [96] and RELICT [97]. In the inflationary model with the non-minimal inflaton field (4.19), the perturbation spectrum compatible with these observations can be obtained for coupling constants of order $\lambda/\xi^2 \sim 10^{-10}$ [86, 98] (experimental restriction on gauge-invariant [99] density perturbation $P_\zeta(k) = N_k^2(\lambda/\xi^2)/8\pi^2$ in the kth mode which crosses the horizon at the moment of expansion $\exp N_k$). The main advantage of this model is the possibility to avoid unnaturally low values of λ in the minimal inflaton model [40, 75] and substitute them by those compatible with the grand unification theory, $\lambda \simeq 0.05$, provided that constant $\xi \simeq -2\times 10^4$ is chosen to be of the same order as the ratio of Planckian to typical grand unification scales, $|\xi| \sim m_{\mathrm{P}}/v$. With such coupling constants, the known bound on the parameter of exponential expansion during the inflationary stage $\exp\big[N(\varphi_I)\big] \geqslant \exp(60)$ when generated by the probability peak (4.32) leads to enormously large anomalous scaling $Z \sim 10^{11}$. A remarkable feature of the non-minimal inflaton model is that such a value may be induced by large constant ξ. Indeed, the expression for $Z_{1\text{-loop}}$, well-known in the generic theory [6, 94], contains the fourth-order contribution in effective particle masses which is easy to calculate on the de Sitter background [100]

$$Z_{1\text{-loop}} = (12H^4)^{-1}\left(\sum_\chi m_\chi^4 + 4\sum_A m_A^4 - 4\sum_\psi m_\psi^4\right) + \ldots, \tag{4.33}$$

where the summation is over all Higgs' scalars χ, vector gauge bosons A, and spinors ψ. Their effective masses for large φ are actually given by contributions $m_\chi^2 = \lambda_\chi\varphi^2/2$, $m_A^2 = g_A^2\varphi^2$, $m_\psi^2 = f_\psi^2\varphi^2$ induced, through the Higgs mechanism, by their Lagrangian of interaction with the inflaton field

$$\boldsymbol{L}_{\text{int}} = \sum_\chi \frac{\lambda_\chi}{4}\chi^2\varphi^2 + \sum_A \frac{1}{2}g_A^2 A_\mu^2\varphi^2 + \sum_\psi f_\psi\varphi\bar\psi\psi + {} + \text{interaction with derivatives}. \tag{4.34}$$

Thus, by virtue of relation $\varphi^2/H^2 = 12|\xi|/\lambda$, full anomalous scaling of the theory

$$Z_{1\text{-loop}} = 6\,\frac{\xi^2}{\lambda}\,\boldsymbol{A} + O(\xi), \tag{4.35}$$

$$\boldsymbol{A} = \frac{1}{2\lambda}\left(\sum_\chi \lambda_\chi^2 + 16\sum_A g_A^4 - 16\sum_\psi f_\psi^4\right) \tag{4.36}$$

contains the same large dimensionless ratio $\xi^2/\lambda \simeq 10^{10}$ and the universal combination of coupling constants $\boldsymbol{A}$ defined by the particle model (neither graviton nor inflaton field contributes to $\boldsymbol{A}$, as well as gravitino when it does not interact with inflaton).

For such $Z_{1\text{-loop}}$, parameters of the probabilistic inflationary peak have the form

$$\varphi_I = m_{\mathrm{P}}\sqrt{\frac{8\pi(1+\delta)}{|\xi|\boldsymbol{A}}}, \qquad \sigma = \frac{\varphi_I}{\sqrt{12\boldsymbol{A}}}\,\frac{\sqrt{\lambda}}{|\xi|}, \tag{4.37}$$

$$H(\varphi_I) = m_{\mathrm{P}}\,\frac{\sqrt{\lambda}}{|\xi|}\sqrt{\frac{2\pi(1+\delta)}{3\boldsymbol{A}^2}}, \qquad N(\varphi_I) = \frac{8\pi^2}{\boldsymbol{A}} \tag{4.38}$$

and satisfy the constraint $\boldsymbol{N}(\varphi_I) \geqslant 60$ with the only condition on $\boldsymbol{A} \leqslant 1.3$.

At $\delta \ll 1$ $(\delta \sim 8\pi/|\xi|$ for $|\xi| \sim m_{\mathrm{P}}/v)$ and $\boldsymbol{A} \simeq 1$, the resultant numerical parameters describe a very sharp and narrow probability peak with the sub-Planckian Hubble constant

$$\varphi_I \simeq 0.03 m_{\mathrm{P}}, \qquad \sigma \simeq 10^{-7} m_{\mathrm{P}}, \qquad H(\varphi_I) \simeq 10^{-5} m_{\mathrm{P}}, \tag{4.39}$$

which forms a realistic region for the inflationary scenario. It is worth noting that the relative width of the peak

$$\frac{\sigma}{\varphi_I} \sim \frac{\Delta H}{H} \sim 10^{-5} \tag{4.40}$$

corresponds to the observed anisotropy level of microwave background even though it is not quite clear whether quantum dispersion σ is directly observable at present, due to stochastic noise of the same order of magnitude generated during inflation and superimposed on σ.

All these conclusions are sufficiently universal and (except the choice of $|\xi|$ and λ) universally depend on parameter $\boldsymbol{A}$ (4.36) of the particle model in the early Universe. This parameter must satisfy the constraints

$$\frac{\lambda}{|\xi|} \ll \boldsymbol{A} \leqslant 1.3, \tag{4.41}$$

in order that the large $|\xi|$ approximation in (4.35) be adequate to guarantee the suppression of over-Planckian energy scales (the lower bound) and sufficiently long inflation (the upper bound). These constraints suggest the (quasi)supersymmetric nature of the model of particles in the early Universe [95] since only supersymmetry is likely to ensure the balance of contributions of bosons and fermions in (4.36) and fit the constant $\boldsymbol{A}$ in the range (4.41).

Let us briefly summarise the results obtained by applying gravitational tunnelling methods in the theory of quantum origin of the early Universe. Evidently, the quantum mechanism which suppresses the contribution of over-Planckian scales also generates a narrow probability peak in the distribution of tunnelling inflationary universes and suggests (quasi)supersymmetric nature of the field model for the early Universe. This peak occurs at a sub-Planckian value of the Hubble constant which justifies the semiclassical expansion for quantum gravitational effects and is in a good agreement with observations of the microwave radiation background in the non-minimal inflaton field model. A distinctive feature of the theory is that its results are based on one small parameter, i.e. the dimensionless ratio of two fundamental scales, that of the grand unification theory and the Planck scale given by the combination of constants $\sqrt{\lambda}/|\xi| \simeq 10^{-5}$. Certainly, a specific value of this combination also needs explanation which may be available with the help of renormalization group methods (or their generalization to perturbatively non-renormalizable theories) [88].

From the viewpoint of quantum cosmology of the early Universe, these results give a strong preference to the tunnelling quantum state as opposed to the Hartle – Hawking state. Advantages of either state have been a matter of a long-standing discussion [40, 75, 60, 84, 101, 78]. Today, the tunnelling state appears to be more useful and conceptually clear in the cosmological context because its interpretation does not require using of questionable ideas of third quantization which arise in the case of the Hartle – Hawking wave function, e.g. its expansion in positive and negative components and separate calculation of their probability distributions. On the other hand, the formulation of the tunnelling state is not aesthetically closed because it uses a boundary condition with positive frequency behind the potential barrier, normalization to a constant under the barrier at $a = 0$, the requirement of normalizability with respect to inhomogeneous modes f, etc. All this is quite different from the closed formulation in the form of functional integral for the Hartle – Hawking wave function which automatically guarantees many of these properties. At the same time, the tunnelling prescription for the wave function seems useless outside the cosmological context, e.g. in physics of wormholes and black holes. Moreover, at the overlap of cosmological problems and the theory of virtual black holes, it leads to the unnatural conclusion that the quantum birth of large black holes is more likely to occur than that of the Planck-size black holes [102].

These arguments can hardly be conclusive beyond the scope of a consistent theory that must establish the status of third quantization and solve an open problem of correct quantization of the conformal mode in Einstein's Lagrangian, etc. This and related problems do not seem to play a major role in the theory of quantum birth of the early Universe. It should be emphasized that the normalizability criterion for the wave function and algorithm (4.25) do not refer to the low energy limit $\varphi \to 0$ where the naively calculated Hartle – Hawking wave function blows up to infinity, and the slow roll approximation becomes inapplicable. Fortunately, this region is separated from the probabilistic inflationary peak by a wide 'desert' with practically zero density of quantum distribution which justifies the above conclusions neglecting the phenomena of ultra-infrared physics of baby universes and Coleman's cosmological constant theory [30].

5. Conclusions

The key issue of the present paper is that the principal object of quantum gravity, the so-called 'wave function of the Universe', has no sense from the quantum-mechanical point of view because it obeys constraint equations and, consequently, cannot be normalized in superspace. The normalizable wave function and its dynamical evolution, the Shrödinger equation, which it obeys are introduced by imposing gauge conditions (3.18) or (3.22), parametrized by number t, on superspace coordinates. The question arises: Does the theory make sense when all the dynamics is given by the choice of gauge conditions, i.e. when "What we call time changes the physics" (Karel Kuchar in Ref. [103], p. 504)?

Note that this is exactly the case in both classical and quantum GTR where 'dynamics' understood as explicit time-dependence of fields obviously changes with transformation of coordinates; in other words, it depends on the choice of reference frame (i.e. the choice of gauge conditions). However, in the classical GTR there is a standard procedure to transform one reference frame to another, whereas in quantum gravity the situation is less transparent. Quantum ADM theories constructed using different gauge conditions may prove to be physically nonequivalent.

It is well-known that the quantum vacuum state of the Minkowski space is not 'empty' in an accelerated reference frame [104], and a freely falling observer who falls in the black hole space can see neither the event horizon nor Hawking radiation even though they are objectively recorded by instruments at rest in the asymptotic region. This 'discrepancy' called 'the information paradox' has recently led to the formulation of a 'principle of black hole complementarity' [105, 106] which, according to the authors, "will change all our ideas of quantum gravity".

We believe, however, that the problem is not so much that the dynamics is given by the choice of gauge conditions; rather, it is in the necessity to formulate the rules for the construction of gauge conditions corresponding to the choice of physical reference frames given by physical bodies, which is not easy to do in the framework of quantum theory: suffice it to recall that even such usual words as 'free fall' has the operator meaning in quantum gravity. In any case, the choice of such *physically motivated* gauge conditions must depend on the dynamical properties of the system and possibly on the background geometry (leading approximation in semiclassical expansion in powers of the Planck constant). Such a dependence of quantization procedure on the initial point of perturbation theory introduces fundamental nonlinearity which is absent in the conventional quantum theory.

The problem of choosing gauge conditions has a different aspect in the third quantization formalism where the 'wave function of the Universe' is taken as the operator. Obviously, the definition of equal-time commutators necessary in this approach requires that the notion of 'synchronization', i.e. slicing of superspace by hypersurfaces, be introduced which is achieved by the choice of gauge conditions. Conversely, the use of the 'S-matrix' language covariant in superspace which does not require gauge fixation implies the introduction of sources into the right-hand side of the Wheeler – DeWitt equation, that is extension beyond the 'mass-shell' of the constraint equation.

It should be emphasized that a direct analogy of quantum gravity with the relativistic particle theory and its second quantization is relevant only as far as minisuperspace models are concerned. In the general case, ∞^3 Wheeler – DeWitt equations (by the number of points in the 3-space), rather than one, are valid. Thus, strictly speaking, quantum gravity corresponds to ∞^3 relativistic particles propagating in a finite-dimensional (by the number of fields estimated with regard for supermomentum constraints) minisuperspace and interacting with one another due to the presence of terms with $\mathbf{x}$ derivatives in the original Lagrangian. Formulas of the present paper are certainly applicable to this general case as well. However, quantization by the ADM reduction method in the problem of two interacting relativistic particles (not to say about their infinite number) is non-trivial and requires special consideration. This is quite different from the case of 'one-particle' problems discussed in Sections 3.5 and 3.7.

Returning to basic principles, we shall try to specify the initial 'technical' reason for quantum gravity to be so difficult and so interesting a science. It is known that the trajectory of a nonrelativistic particle propagating in a stationary potential $U(\mathbf{x})$ with energy E can be obtained by varying the so-called 'truncated' Maupertuis action [107]

$$S_{\mathrm{M}} = \int \sqrt{T(E-U)}\,\mathrm{d}t \tag{5.1}$$

(T is kinetic energy), which formally describes the propagation of a relativistic particle with variable mass or, in other words, its propagation in a conformally-flat space with the scale factor $(E-U)$. Classical extremals (with energy E) of the standard action with the Lagrangian $(T-U)$ and action (5.1) obviously coincide, but quantization of the theory (5.1) , i.e. a system placed on the constant energy surface, leads to essentially different results which are impossible to obtain in the framework of a non-relativistic problem. (Quantization of systems with constraints has been described in Ref. [108]). In GTR, an analog of (5.1) was obtained by substituting the g_{00}-component of the metric expressed from the (00)-component of Einstein's equations into the Einstein action; it was first suggested by Baierlein, Sharp, Wheeler (BSW) in 1962 [109] (see also Refs [110, 111]).

Generally speaking, BSW gravitation (i.e. gravitation in which the Wheeler – DeWitt constraint is valid) in asymptotically flat spaces is not equivalent to standard GTR and can be obtained from GTR by imposing an additional constraint that the surface integral which defines the energy-momentum of gravitational field and matter vanishes. Indeed, most of the conceptual problems of quantum gravity (the problem of time, etc.) result from imposing this constraint which is always valid in closed spaces. The difficulties encountered in the interpretation of quantum gravitation near the 'turning point' are due to degeneracy of the action similar to (5.1) at $E-U=0$. In connection with this it appears appropriate to mention the time arrow problem which is not considered in the present paper but was at the focus of A D Sakharov's interests (see the introductory article in the jubilee issue of *Uspekhi Fizicheskikh Nauk* (*Soviet Physics-Uspekhi*) [112]). In an early paper, Hawking argued [113] that maximum expansion of the Universe must be accompanied by the reversal of thermodynamic arrow of time, but later he regarded this assertion to be his most serious mistake. (See Refs [114, 115] and also a recent publication [116] in which the authors still advocate the idea of reversal of the direction of time following the maximum expansion of the Universe). There is little doubt that these conflicting views reflect real difficulties of the theory.

Another fundamental issue is related to the problem of a 'turning point' in the vicinity of which the semiclassical approximation is inapplicable. The question is whether quantum gravity with the superhamiltonian constraint admits the construction of a semiclassical wave packet describing stationary spacetime, e.g. the Minkowski space. In a sense, the stationary state is a 'permanent turning point'. In view of this, many authors maintain that GTR with the Wheeler – DeWitt constraint satisfies Mach's principle which treats the empty Minkowski space as a physically inadmissible object [23, 24, 103, 110, 117].

In conclusion, it is worthwhile to dwell on two virtually incompatible approaches to explaining fundamental properties of the observable Universe: dimensionality, signature, constants ... One, known as the 'anthropological selection' principle, was advocated by A D Sakharov: ''We believe, in the spirit of the anthropological principle, that the observable Universe is singled out by the sum of parameter values favouring the development of life and intellect. In particular, the signature ... may be one of such parameters'' [2]. Moreover: ''In space P, an infinite number of U-insertions should be considered (for the whole complex of trajectories and even for a single trajectory); parameters of the infinite number of them may be arbitrarily close to the parameters of the observable Universe. It is therefore supposed that the number of universes similar to our Universe, in which structures, life and intellect are possible, may be infinite. This does not rule out the possibility that life and intellect occur in an infinite number of universes of essentially different classes, which form either finite or infinite number of classes of 'similar' universes including those with the signature different from the signature of our Universe'' [2]. (Here, P is Euclidean space of positive signature, U is the Universe with single time). Sakharov was adherent to the idea of multiple universes which naturally arises in third quantization of gravity and also in chaotic inflation models (see many of his works and also the Nobel lecture).

This incredible variety of options makes one think of hiding oneself in the shell of naive anthropomorphism. The 'dynamical predetermination' approach opposite to the anthropic principle does not deny the existence of 'spacetime foam' on the Planck scale, but it relies on the possibility to formulate principles and equations which would allow to calculate fundamental constants and explain why universes with dimensionalities and signatures other than $3+1$ can not be large and semiclassical. An infinite number of universes postulated by Sakharov and all talks about their spacetime characteristics make sense only in the semiclassical approximation; the situation may be much less ambiguous at the fundamental quantum level. Motivated by basic principles strict requirements for existence of the semiclassical limit (a phase of broken symmetry of general-coordinate diffeomorphisms) may be of predictive value for explaining properties of a macroscopic universe. Here, quantum gravity clears the way for generating new ideas. However, the most important, 'crazy' (as Bohr put it) idea remains to be suggested. Having no claim on throwing light on all the particulars of the problem of 'dynamical predetermination', which needs a special review to be comprehensively discussed, we would like to bring to readers' attention Refs [118 – 120] in which the authors endeavoured to explain 3-dimensionality of space in the framework of string theory. The theory of virtual wormholes and 'big fix' of Coleman, Giddings, and Strominger [30] (see Section 1.6 of the Introduction) appears to have been designed in an attempt to realize the 'dynamical predetermination' principle. Unfortunately, it has not yet proved its value, despite the original optimism of its supporters. Results presented in Section 4, which describe the distribution function with a maximum for the Hubble constant determining the inflation rate, should also be considered in the context of striving for dynamic (as opposed to anthropic) predetermination of the properties of the observable Universe.

Acknowledgements
The authors would like to thank K Gjotterud, A Yu Kamenshchik, V A Rubakov, and I V Tyutin for helpful discussions.

We are grateful to the Russian Foundation for Basic Research for the support of this work under grant 96-02-16287-a.

One of the authors (AOB) is also grateful to the Russian Foundation for Basic Research for support under grants 96-02-16295 and 96-01-00482, to the International Science Foundation (Soros Foundation) and the Russian Government (grant MQY300), and for the grant of the European Community (grant INTAS-93-493). This work was also made possible in part due to the support by the Russian Research Project 'Kosmomikrofizika (Cosmomicrophysics)'.

References

1. *Akademik A D Sakharov. Nauchnye Trudy* (Academician A D Sakharov. Collected Works) (Eds B L Al'tshuler, L V Keldysh (chairman), D A Kirzhnits, V I Ritus) (Moscow: Tsentr-Kom OTF FIAN, 1995)
2. Sakharov A D *Zh. Eksp. Teor. Fiz.* **87** 375 (1984) [*Sov. Phys. JETP* **60** 214 (1984)]
3. *Usp. Fiz. Nauk* **161** (5) (1991) [*Sov. Phys. Usp.* **34** (5) (1991)]
4. Sakharov A D *Gor'kiĭ, Moskva, dalee vezde* (Gor'kiĭ, Moscow, then everywhere) (New York: Izd. im Chekhova, 1990); *Znamya* (9 – 10) (1991)
5. Dirac P A M *Lektsii po Kvantovoĭ Mekhanike* (Lectures on Quantum Mechanics) (New York: Yeshiva University, 1966) [Translated into Russian (Moscow: Mir, 1968)]
6. DeWitt B S *Phys. Rev.* **160** 1113 (1967); **162** 1195 (1967)
7. Wheeler J A, in *Batelles Recontres* (Eds C DeWitt, J A Wheeler) (New York: Benjamin, 1968)
8. Barvinsky À *Phys. Rep.* **230** 237 (1993)
9. Arnowitt R, Deser S, Misner C, in *Gravitation: an Introduction to Current Research* (Ed. L Witten) (New York: Wiley, 1962)
10. Fradkin E S, Vilkovisky G A *Phys. Lett. B* **55** 224 (1975); CERN Report TH-2332 (1977); Batalin I A, Vilkovisky G A *Phys. Lett. B* **69** 309 (1977)
11. Becchi C, Rouet A, Stora R *Comm. Math. Phys.* **42** 127 (1975); Tyutin I V, Preprint No. 39 FIAN (1975); Kugo T, Ojima I *Suppl. Progr. Theor. Phys.* **66** 1 (1979)
12. Hawking S W, in *General Relativity. An Einstein Centenary Survey* (Eds S W Hawking, W Israel) (Cambridge, London, New York, New Rochelle, Melbourne, Sydney: Cambridge Univ. Press, 1979) p. 747
13. Barvinsky A O, in *Proc. of GR9 540* (*Jena, 1980*)
14. Barvinsky A O, Ponomariov V N *Phys. Lett. B* **167** 287 (1986)
15. Teitelboim C *Phys. Rev. D* **25** 3159 (1982); **28** 297, 310 (1983)
16. Halliwell J J *Phys. Rev. D* **38** 2468 (1988)
17. Hartle J B, Hawking S W *Phys. Rev. D* **28** 2960 (1983)
18. Hawking S W *Nucl. Phys. B* **239** 257 (1984)
19. DeWitt B S, in *Quantum Field Theory and Quantum Statistics* (Eds I A Batalin, C J Isham, G V Vilkovisky) (Bristol: Adam Hilger, 1987) p. 191
20. Vilkovisky G A *Nucl. Phys. B* **234** 125 (1984); *Class. Quant. Grav.* **9** 895 (1992); in *Quantum Theory of Gravity* (Ed. S M Christensen) (Bristol: Adam Hilger, 1984)
21. Epp R J, Kunstatter G, Toms D J *Phys. Rev. D* **47** 2474 (1993)
22. Barbero J F, Perez-Mercader J *Phys. Rev. D* **48** 3663 (1993)
23. Al'tshuler B L, Boyarsky A M, Neronov A Yu "The problem of classical limit in quantum cosmology: the effective action language", in *Report at the Seminar "Quantum Gravity. In Memory of Academician Moiseĭ Markov", June 12 – 19, Moscow, 1995,* gr-qc 9511069 (to be published in World Scientific)
24. Altshuler B L, Boyarsky A M, Neronov A Yu *Gravitation and Cosmology* **1** 191 (1995)
25. Isham C J *J. Math. Phys.* **35** 6360 (1994)
26. Barvinsky A O, Kamenshchik A Yu *Phys. Rev. D* **52** 743 (1995)
27. Wheeler J A *Ann. Phys.* (*N.Y.*) **2** 604 (1957); in *Geometrodynamics* (New York: Academic, 1962)
28. Banks T *Nucl. Phys. B* **309** 493 (1988)
29. Lavrelashvili G V, Rubakov V A, Tinyakov P G *Pis'ma Zh. Eksp. Teor Fiz.* **46** 167 (1987) [*JETP Lett.* **46** 167 (1987)]; *Nucl. Phys. B* **299** 757 (1988); *Mod. Phys. Lett. A* **3** 1231 (1988)
30. Giddings S, Strominger A *Nucl. Phys. B* **306** 890 (1988); **321** 481 (1988); Coleman S *Nucl. Phys. B* **307** 867 (1988); **310** 643 (1988)
31. Hawking S W, Laflamme R *Phys. Lett. B* **209** 39 (1988); Hawking S W *Phys. Rev. D* **37** 904 (1988); *Nucl. Phys. B* **335** 155 (1990)
32. Rubakov V A *Nucl. Phys. B* **453** 395 (1995)
33. Hawking S W *Nucl. Phys. B* **363** 117 (1991)
34. Bardakci K *Nucl. Phys. B* **271** 561 (1986)
35. Tseytlin A A *Phys. Lett. B* **251** 530 (1990)
36. Bleyer U et al. *Nucl. Phys. B* **429** 177 (1994)
37. Redmount I H, Suen W-M *Phys. Rev. D* **49** 5199 (1994)
38. Frolov V P, Novikov I D *Phys. Rev. D* **42** 1057 (1990)
39. Frolov V P, in *Sakharov Memorial Lectures in Physics* (Eds L V Keldysh, V Ya Fainberg) (New York: Nova Science Publ. Inc., 1992) p. 1067
40. Linde A D *Particle Physics and Inflationary Cosmology* (Contemporary Concepts in Physics Vol. 5) (Chur, Switzerland: Harwood Academic, 1990)
41. Greensite J *Phys. Lett. B* **300** 34 (1993)
42. Carlini A, Greensite J *Phys. Rev. D* **49** 866 (1994)
43. Kriele M, Martin J *Class. Quant. Grav.* **12** 503 (1995)
44. Embacher F *Gravitation and Cosmology* **1** 46 (1995)
45. Hellaby C, Dray T *Phys. Rev. D* **49** 5096 (1994)
46. Blencowe M P, Duff M J *Nucl. Phys. B* **310** 387 (1988)
47. Sakharov A D *Dokl. Akad. Nauk SSSR* **177** 70 (1967); *Teor. Mat. Fiz.* **23** 178 (1975)
48. Green M B, Schwarz J H, Witten E *Superstring Theory* (Cambridge: Cambridge Univ. Press, 1987)
49. Polyakov A M *Gauge Fields and Strings* (Chur, Switzerland: Harwood Academic Publishers, 1987)
50. *J. Math. Phys.* **36** 11 (1995) Special Issue on Quantum Gravity and Diffeomorphism Invariant Quantum Field Theory
51. Halliwell J J *Int. J. Mod. Phys. A* **5** 2473 (1990)
52. Isham C J *Structural Issues in Quantum Gravity* (Plenary Session lecture given at the GR 14 Conference, Florence, August 1995)
53. *Sakharov Memorial Lectures in Physics. Proc. of the First International Sakharov Conference on Physics, Moscow, 21 – 31 May, 1991* (Eds L V Keldysh, V Ya Fainberg) (New York: Nova Science Publ. Inc., 1992)
54. Kuchar K *J. Math. Phys.* **17** (1976) 777
55. Faddeev L D *Teor. Mat. Fiz.* **1** 3 (1969)
56. Barvinsky A O, Krykhtin V *Class. Quant. Grav.* **10** 1957 (1993)
57. Kuchar K, in *Proc. of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics* (Eds G Kunstatter, D Vincent, J Williams) (Singapore: World Scientific, 1992)
58. Geroch R P *J. Math. Phys.* **8** 782 (1967)
59. DeWitt B, in *Proc. of the Third Seminar on Quantum Gravity* (Eds M A Markov, V A Berezin, V P Frolov) (Singapore: World Scientific, 1985)
60. Vilenkin A *Phys. Lett. B* **117** 25 (1982); *Phys. Rev. D* **27** 2848 (1983)
61. Lapchinsky V G, Rubakov V A *Acta Phys. Polon. B* **10** 1041 (1979)
62. Banks T *Nucl. Phys. B* **249** 332 (1985)
63. Kiefer C, Preprint Freiburg THEP-95/1, gr-qc/9501001
64. Batalin I A, Fradkin E S *Riv. Nuovo Cimento* **9** 1 (1986); *Ann. Inst. Henri Poincare* **49** 145 (1988)
65. Barvinsky A O *Phys. Lett. B* **241** 201 (1990)
66. Barvinsky A O *Class. Quant. Grav.* **10** 1985 (1993)
67. DeWitt-Morette C *Ann. Phys.* **97** 367 (1976)
68. Gribov V N *Nucl. Phys. B* **139** 1 (1978)
69. Barvinsky A O, Kamenshchik A Yu *Phys. Rev. D* **50** 5093 (1994)
70. Halliwell J J, Hartle J B *Phys. Rev. D* **41** 1815 (1990)
71. Gibbons G W, Hartle J B *Phys. Rev. D* **42** 2458 (1990)

72. York J W (Jr) *Phys. Rev. Lett.* **26** 1656 (1971); **28** 1082 (1972)
73. O'Murchadha N, York J W (Jr) *J. Math. Phys.* **14** 1551 (1973)
74. Ashtekar A *Non-Perturbative Quantum Gravity* (Singapore: World Scientific, 1991)
75. Linde A D *Phys. Lett. B* **129** 177 (1983)
76. Sakharov A D *Zh. Eksp. Teor. Fiz.* **49** 345 (1965) [*Sov Phys. JETP* **22** 241 (1966)]
77. Hawking S W, Page D N *Nucl. Phys. B* **264** 185 (1986)
78. Vilenkin A *Phys. Rev. D* **37** 888 (1988)
79. Barvinsky A O, Kamenshchik A Yu *Class. Quant. Grav.* **7** L181 (1990)
80. Barvinsky A O, Kamenshchik A Yu *Phys. Lett. B* **332** 270 (1994)
81. Mottola E, Lapedes A *Phys. Rev. D* **27** 2285 (1983)
82. Laflamme R *Phys. Lett. B* **198** 156 (1987)
83. Barvinsky A O, Kamenshchik A Yu, Mishakov I V "Quantum Creation of the Early Universe", to be published in *Nucl. Phys. B*
84. Linde A D *Zh. Eksp. Teor. Fiz.* **87** 369 (1984) [*Sov. Phys. JETP* **60** 211 (1984)]; *Lett. Nuovo Cimento* **39** 401 (1984); Zel'dovich Ya B, Starobinskiĭ A A *Sov. Astron. Lett.* **10** 135 (1984); Vilenkin A *Phys. Rev. D* **30** 509 (1984)
85. Vachaspati T, Vilenkin A *Phys. Rev. D* **37** 898 (1988)
86. Salopek D S, Bond J R, Bardeen J M *Phys. Rev. D* **40** 1753 (1989)
87. Page Don N *J. Math. Phys.* **32** 3427 (1991)
88. Barvinsky A O, Kamenshchik A Yu, Karmazin I P *Phys. Rev. D* **48** 3677 (1993)
89. Spokoiny B L *Phys. Lett. B* **147** 39 (1984); Fakir R, Unruh W G *Phys. Rev. D* **41** 1783 (1990)
90. Halliwell J J, Hawking S W *Phys. Rev. D* **31** 1777 (1985)
91. Wada S *Nucl. Phys. B* **276** 729 (1986)
92. Barvinsky A O, Kamenshchik A Yu "Quantum Origin of the Early Universe and the Energy Scale of Inflation", in *Proc. VI Sem. "Quantum Gravity" Moscow, 1995*, gr-qc/9510032
93. Barvinsky A O *Phys. Rev. D* **50** 5115 (1994)
94. 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)
95. Kamenshchik A Yu *Phys. Lett. B* **316** 45 (1993)
96. Smoot J et al. *Astrophys. J.* **396** L1 (1992)
97. Strukov I et al. *Pis'ma Astron. Zh.* **18** 387 (1992); *Month. Not. Roy. Astron. Soc.* **258** 37p (1992)
98. Salopek D S, Preprint DAMTP R92/41 (1992)
99. Bardeen J M, Steinhardt P J, Turner M S *Phys. Rev. D* **28** 679 (1983)
100. Allen B *Nucl. Phys. B* **226** 228 (1983); Fradkin E S, Tseytlin A A *Nucl. Phys. B* **234** 472 (1984)
101. Rubakov V A *Phys. Lett. B* **148** 280 (1984)
102. Bousso R and Hawking S W, Preprint DAMTP/R-95/33, gr-qc/9506047
103. *Mach's Principle: from Newton's Bucket to Quantum Gravity* (Eds J Barbour, H Pfister) *Einstein Studies* Vol. 6 (Boston, Basel, Berlin: Birkhauser, 1995)
104. Ginzburg V L, Frolov V P *Trudy FIAN im P N Lebedeva* **197** 8 (1989)
105. Susskind L, Thorlacius L, Uglum J *Phys. Rev. D* **48** 3743 (1993)
106. Susskind L, Thorlacius L *Phys. Rev. D* **49** 966 (1994)
107. Landau L D, Livshits E M *Mekhanika* (Mechanics) (Moscow: Fizmatgiz, 1958)
108. Gitman D M, Tyutin I V *Kanonicheskoe Kvantovanie poleĭ so svyazyami* (Canonical Quantization of Fields with Constraints) (Berlin, New York: Springr-Verlag, 1990) [Translated into Russian (Moscow: Nauka, 1986)]
109. Baierlein R F, Sharp D H, Wheeler J A *Phys. Rev.* **126** 1864 (1962)
110. Barbour J B *Class. Quant. Grav.* **11** 2853, 2875 (1994)
111. Greensite J *Phys. Rev. D* **49** 930 (1994)
112. Al'tshuler B L See Ref. [3], p. 3
113. Hawking S W *Phys. Rev. D* **32** 2489 (1985)
114. Page Don N, in Ref. [53], p. 1031
115. Gell-Mann M, Hartle J B, in Ref. [53], p. 1151
116. Kiefer C, Zeh H D *Phys. Rev. D* **51** 4145 (1995)
117. Lynden-Bell D, Katz J, Bicak J *Month. Not. Roy. Astr. Soc.* **272** 150 (1995)
118. Cleaver G B, Rosenthal P J *Nucl. Phys. B* **457** 621 (1995)
119. Brandenberger R, Vafa C *Nucl. Phys. B* **316** 391 (1989)
120. Tseytlin A A, Vafa C *Nucl. Phys. B* **372** 443 (1992)