Gravitation & Cosmology, Vol. 10 (2004), No. 4 (40), pp. 257–268

REVIEW ARTICLE

THE PROBLEM OF TIME AND GAUGE INVARIANCE IN THE QUANTIZATION OF COSMOLOGICAL MODELS. II. RECENT DEVELOPMENTS IN THE PATH INTEGRAL APPROACH

T.P. Shestakova[1]† and C. Simeone[2]‡

† *Department of Theoretical and Computational Physics, Rostov State University, Sorge Str. 5, Rostov-on-Don 344090, Russia*

‡ *Departamento de Física, Facultad de Ciencias Exactas y Naturales Universidad de Buenos Aires, Ciudad Universitaria Pabellón I - 1428, Buenos Aires, Argentina and Instituto de Astronomía y Física del Espacio C.C. 67, Sucursal 28 - 1428 Buenos Aires, Argentina*

Received 4 June 2004

This is the second part of the work devoted to the problem of time in quantum cosmology. Here we consider in detail two approaches within the scope of Feynman path integration scheme: The first, by Simeone and collaborators, is gauge-invariant and lies within the unitary approach to a consistent quantization of gravity. It is essentially based on the idea of deparametrization (reduction to physical degrees of freedom) as a first step before quantization. The other approach by Savchenko, Shestakova and Vereshkov is rather radical. It is an attempt to take into account the peculiarities of the Universe as a system without asymptotic states, which leads to the conclusion that quantum geometrodynamics constructed for such a system is, in general, a gauge-noninvariant theory. However, this theory is shown to be mathematically consistent, and the problem of time is solved thereby in a natural way.

Проблема времени и калибровочная инвариантность в квантовании космологических моделей. II. Новые результаты в формализме интегрирования по путям
Т.П. Шестакова, К. Симеоне

Статья является второй частью работы, посвященной проблеме времени в квантовой космологии. Здесь мы подробно рассматриваем два подхода в рамках фейнмановского интегрирования по траекториям. Первый, предложенный Симеоне и его коллегами, калибровочно-инвариантен и лежит в русле унитарного подхода к непротиворечивому квантованию гравитации. Он существенно опирается на идею депараметризации (редукции к физическим степеням свободы) как предварительного этапа перед квантованием. Другой подход, предложенный Верешковым, Савченко и Шестаковой, является в достаточной мере радикальным. Это попытка принять во внимание особенности Вселенной как системы без асимптотических состояний, что приводит к заключению, что квантовая геометродинамика, построенная для такой системы, представляет собой, вообще говоря, калибровочно-неинвариантную теорию. Показано, однако, что эта теория математически непротиворечива, а проблема времени решается в ней естественным образом.

1. Introduction

In Part I of our work [39], we have considered the most representative approaches to the well-known problem of time in quantum cosmology which lie in the framework of canonical quantization. Unfortunately, most of these proposals can be only applied to restricted classes of models. The most interesting and promising approaches which go beyond the minisuperspace approximation are the proposals by Brown and Kuchař [8] and by Barvinsky and collaborators [1–4], the latter having been formulated mainly in the framework of the Feynman path integral formalism.

The main object of the path integral approach [12, 16, 17] is a transition amplitude between two states which is obtained as a sum over all histories of the exponential of the action. For a constrained system, divergences yielded by overcounting of paths in phase space, which are physically equivalent, should be avoided by imposing gauge conditions which select one path from each equivalence class [13, 14]. In its phase space form, the propagator then reads

$$\langle q_2^i | q_1^i \rangle = \int Dq^i Dp_i DN \delta(\chi) |[\chi, \mathcal{H}]| \exp\left(iS[q^i, p_i, N]\right). \quad (1)$$

Here $\chi = 0$ is a gauge-fixing function, and $|[\chi, \mathcal{H}]|$ is the Faddeev–Popov determinant which makes the result

[1] e-mail: shestakova@phys.rsu.ru
[2] e-mail: csimeone@df.uba.ar

independent of the gauge choice. Admissible gauge conditions are those which can be reached from any path by performing a gauge transformation compatible with the symmetries of the action.

Since, in gravitational dynamics, the Hamiltonian generates the evolution and also acts as a generator of gauge transformations, it is natural to think that time could be defined by means of gauge fixing, so that the resulting non-divergent amplitudes would include a clear notion of evolution. But the problem arises that the gauges defining time in terms of the canonical variables are the so-called canonical gauges, which can be imposed only if the constraints are linear in the canonical momenta, while the Hamiltonian constraint in the gravitational action is quadratic in momenta [46]. This seemed to be an obstacle for a programme of deparametrization based on this idea; however, we will show in Sec. 2 that, for a class of cosmological models, this can be solved by associating to them an ordinary gauge system (that is, a system with constraints linear in momenta), so that gauge fixing in the gauge system defines time for the corresponding minisuperspace [41].

In Sec. 2, we shall follow Simeone and collaborators [11, 15, 20–22, 40–44]. The connection between fixing an admissible gauge condition and the definition of time will be considered in detail in Sec. 2.1. In Sec. 2.2, we will describe a special canonical transformation that gives rise to an action for a system with a zero Hamiltonian and a constraint which is linear in momenta. On this way, we face the problem of observations mentioned in Part I of our work: new canonical variables appear to be conserved quantities since they commute with a new Hamiltonian. So we need another canonical transformation which leads to a time-dependent Hamiltonian. This will be discussed in Sec. 2.3. We will arrive at a formulation in terms of true degrees of freedom in what we call the reduced phase space. It allows us to define a transition amplitude through a path integral by the usual Faddeev–Popov procedure in Sec. 2.4. Examples will be given in Sec. 2.5.

In the approaches of Barvinsky and of Simeone and collaborators, time is introduced into the theory by means of a time-dependent gauge condition. In Sec. 3 it will be shown that time may appear as a consequence of breaking down the gauge invariance of the theory, even if the gauge condition is time-independent. In the approach presented in the papers by Savchenko, Shestakova and Vereshkov [32–36], the authors argued that the breakdown of gauge invariance is inevitable since the Universe as a physical system does not possess asymptotic states. This prevents imposing asymptotic boundary conditions which eventually ensure the gauge invariance. This will be discussed in Sec. 3.1. In Sec. 3.2, the dynamics of a simple minisuperspace model in extended phase space will be constructed, and its quantum version will be explored in Sec. 3.3. Finally, in Sec. 3.4 we shall touch upon an intriguing question of whether the irreversibility of time could be related to a nontrivial topology of the Universe.

2. Path integral quantization of minisuperspaces as ordinary gauge systems

In this section we shall review our procedure for associating an ordinary gauge system to a minisuperspace model, which allows one to effectively deparametrize the minisuperspace and to obtain a consistent path integral quantization. The analogy between gauge transformations and dynamical evolution, reflected in the equations

$$\frac{dq^i}{d\tau} = N_\mu[q^i, \mathcal{H}^\mu], \qquad \frac{dp_i}{d\tau} = N_\mu[p_i, \mathcal{H}^\mu] \tag{2}$$

and

$$\begin{aligned} \delta_\epsilon q^i &= \epsilon_\mu(\tau)[q^i, \mathcal{H}^\mu], \\ \delta_\epsilon p_i &= \epsilon_\mu(\tau)[p_i, \mathcal{H}^\mu], \\ \delta_\epsilon N_\mu &= \frac{\partial \epsilon_\mu(\tau)}{\partial \tau} - u^{\nu\rho}_\mu \epsilon_\rho N_\nu \end{aligned} \tag{3}$$

is the basic idea leading to the reduction procedure identifying the physical degrees of freedom and time. However, because of the lack of gauge invariance at the end points in the action of gravitation resulting from the quadratic form of the Hamiltonian constraint, admissible gauges do not possess the canonical form $\chi(q^i, p_i, \tau) = 0$; hence, in order to perform the deparametrization, we shall introduce a reformulation of the theory leading to a globally gauge-invariant action [41].

2.1. Gauge fixing and deparametrization

Admissible gauge conditions are those which can be reached from any path by means of gauge transformations, leaving the action unchanged, and such that only one point of each orbit is on the manifold defined by $\chi = 0$. This requires analyzing the possibility of the *Gribov problem* [23, 29], namely, that, depending of the form of the orbits and on the topology of the constraint surface, it may be difficult to intersect it with a gauge condition which is crossed by each orbit only once.

If it is possible to perform a canonical transformation $(q^i, p_i) \to (Q^i, P_i)$ such that the Hamiltonian $\mathcal{H}$ is matched to one of the new momenta, in terms of the new variables the action functional will include a constraint which is linear and homogeneous in momenta. This is equivalent to saying that the canonical variables (Q^i, P_i) describe an ordinary gauge system, so that the canonical gauges $\chi(Q^i, P_i, \tau) = 0$ are admissible.

The condition that a gauge transformation moves a point of an orbit off the surface $\chi = 0$ is fulfilled if

$$[\chi, \mathcal{H}] \neq 0. \tag{4}$$

Now, since Q^0 and P_0 are conjugated variables,

$$[Q^0, P_0] = 1 \tag{5}$$

and if we identify $\mathcal{H} \equiv P_0$, then a gauge condition of the form

$$\chi \equiv Q^0 - T(\tau) = 0 \tag{6}$$

with T a monotonous function is a good choice. Eq. (4) only ensures that the orbits are not tangent to the surface $\chi = 0$; however, since (6) defines a plane $Q^0 =$ const for each τ, if at any τ any orbit was intersected more than once (thus yielding Gribov copies) at another τ, it should be $[\chi, P_0] = 0$. Therefore this gauge fixing procedure avoids the Gribov problem [40].

A connection with the identification of time is as follows: as we have already seen, for a parametrized system whose canonical variables are (q^i, p_i), a global phase time $t(q^i, p_i)$ is a function satisfying the condition [24]

$$[t, \mathcal{H}] > 0. \tag{7}$$

Since the Poisson bracket is invariant under a canonical transformation, from (5) and (7) it follows that a global phase time can be defined for a minisuperspace by imposing on its associated gauge system a gauge condition in terms of the coordinate Q^0. In other words, *a gauge choice for the gauge system defines a particular foliation of spacetime for the corresponding cosmological model* [41]. If a gauge choice avoiding the Gribov ambiguity can be found, then a definition of time which is good everywhere is obtained. A transformation such that $\mathcal{H} = P_0$ can always be found locally; in the next paragraphs we shall show how a canonical transformation, which works in the whole phase space, can be found.

2.2. Gauge-invariant action for a minisuperspace

Here we shall review our procedure of obtaining a gauge-invariant action for cosmological models whose Hamiltonian constraint is such that a solution for its associated $\tau-$independent Hamilton–Jacobi equation can be found. Consider a complete solution [31] $W(q^i, \alpha_\mu, E)$ of the Hamilton–Jacobi equation

$$H\left(q^i, \frac{\partial W}{\partial q^i}\right) = E, \tag{8}$$

where H is not necessarily the original Hamiltonian constraint, but it can be a scaled Hamiltonian, that is, $H = \mathcal{F}^{-1}\mathcal{H}$ with $\mathcal{F}$ a positive-definite function of q^i. If E and the integration constants α_μ are matched to the new momenta $\overline{P}_0$ and $\overline{P}_\mu$, respectively, then $W(q^i, \overline{P}_i)$ turns out to be the generating function of a canonical transformation $(q^i, p_i) \to (\overline{Q}^i, \overline{P}_i)$ defined by the equations

$$p_i = \frac{\partial W}{\partial q^i}, \qquad \overline{Q}^i = \frac{\partial W}{\partial \overline{P}_i}, \qquad \overline{K} = N\overline{P}_0 = NH, \tag{9}$$

where $\overline{K}$ is a new Hamiltonian. The new coordinates and momenta satisfy the conditions

$$[\overline{Q}^\mu, \overline{P}_0] = [\overline{Q}^\mu, H] = 0,$$
$$[\overline{P}_\mu, \overline{P}_0] = [\overline{P}_\mu, H] = 0,$$
$$[\overline{Q}^0, \overline{P}_0] = [\overline{Q}^0, H] = 1.$$

The variables $(\overline{Q}^\mu, \overline{P}_\mu)$ are then *observables*: they commute with the constraint, so that they are gauge-invariant. The resulting action

$$\overline{S}[\overline{Q}^i, \overline{P}_i, N] = \int_{\tau_1}^{\tau_2} \left(\overline{P}_i \frac{d\overline{Q}^i}{d\tau} - N\overline{P}_0\right) d\tau \tag{10}$$

describes a system with a zero true Hamiltonian and a constraint which is linear and homogeneous in momenta (hence canonical gauges would be admissible in a path integral with this action). The action $\overline{S}$ is related to S by

$$\overline{S}[q^i, p_i, N] = \int_{\tau_1}^{\tau_2} \left(p_i \frac{dq^i}{d\tau} - NH\right) d\tau + \left[\overline{Q}^i(q^i, p_i)\overline{P}_i(q^i, p_i) - W(q^i, \overline{P}_i)\right]_{\tau_1}^{\tau_2}, \tag{11}$$

so that the gauge-invariant action $\overline{S}$ differs from the original action S in end-point terms [28]. These terms do not modify the dynamics since they can be included in the action integral as a total derivative with respect to the parameter τ.

2.3. Time and the true degrees of freedom

The observables $\overline{Q}^\mu$ and $\overline{P}_\mu$ are conserved quantities because they commute with $\overline{K} = N\overline{P}_0$. This makes impossible the characterization of the dynamical trajectories of the system by an arbitrary choice of $\overline{Q}^\mu$ at the end points τ_1 and τ_2. To obtain a set of observables such that the choice of the new coordinates is sufficient for characterizing the dynamical evolution, non-conserved variables must be defined, and a new $\tau-$dependent transformation leading to a non-null Hamiltonian must be introduced.

Let us consider the canonical transformation generated by

$$F(\overline{Q}^i, P_i, \tau) = P_0\overline{Q}^0 + f(\overline{Q}^\mu, P_\mu, \tau), \tag{12}$$

which leads to

$$\begin{aligned}
\overline{P}_0 &= \frac{\partial F}{\partial \overline{Q}^0} = P_0 = H \\
\overline{P}_\mu &= \frac{\partial F}{\partial \overline{Q}^\mu} = \frac{\partial f}{\partial \overline{Q}^\mu}, \\
Q^0 &= \frac{\partial F}{\partial P_0} = \overline{Q}^0, \\
Q^\mu &= \frac{\partial F}{\partial P_\mu} = \frac{\partial f}{\partial P_\mu}.
\end{aligned} \tag{13}$$

The generator f defines a canonical transformation in what we call the reduced phase space, which corresponds to the true degrees of freedom of the theory.

The coordinates and momenta (Q^μ, P_μ) are observables because

$$[Q^\mu, P_0] = [P_\mu, P_0] = 0,$$

but they are not conserved quantities because their evolution is determined by the non-zero Hamiltonian

$$K = NP_0 + \frac{\partial f}{\partial \tau} = NH + \frac{\partial f}{\partial \tau}. \qquad (14)$$

Indeed,

$$\begin{aligned} \frac{dQ^\mu}{d\tau} &= \frac{\partial K}{\partial P_\mu} = \frac{\partial^2}{\partial\tau\partial P_\mu} f(\overline{Q}^\mu(Q^\mu, P_\mu), P_\mu, \tau), \\ \frac{dP_\mu}{d\tau} &= -\frac{\partial K}{\partial Q^\mu} = -\frac{\partial^2}{\partial\tau\partial Q^\mu} f(\overline{Q}^\mu(Q^\mu, P_\mu), P_\mu, \tau), \end{aligned} \qquad (15)$$

so that

$$h(Q^\mu, P_\mu, \tau) \equiv \frac{\partial}{\partial \tau} f(\overline{Q}^\mu(Q^\mu, P_\mu), P_\mu, \tau) \qquad (16)$$

is a true Hamiltonian for the reduced system (below we shall give a prescription for choosing f). For the coordinate conjugated to P_0 we have

$$\frac{dQ^0}{d\tau} = [Q^0, K] = N[Q^0, P_0] = N. \qquad (17)$$

The transformation $(\overline{Q}^i, \overline{P}_i) \to (Q^i, P_i)$ yields additional end point terms of the form

$$\Big[Q^\mu P_\mu - f(\overline{Q}^\mu(Q^\mu, P_\mu), P_\mu, \tau)\Big]_{\tau_1}^{\tau_2}.$$

The gauge-invariant action resulting from the two successive canonical transformations $(q^i, p_i) \to (\overline{Q}^i, \overline{P}_i) \to (Q^i, P_i)$ is

$$\mathcal{S}[Q^i, P_i, N] = \int_{\tau_1}^{\tau_2} \left(P_i \frac{dQ^i}{d\tau} - NP_0 - \frac{\partial f}{\partial \tau} \right) d\tau, \qquad (18)$$

and, in terms of the original variables, it includes end point terms:

$$\begin{aligned} \mathcal{S}[q^i, p_i, N] &= \int_{\tau_1}^{\tau_2} \left(p_i \frac{dq^i}{d\tau} - NH \right) d\tau \\ &+ \Big[\overline{Q}^i \overline{P}_i - W(q^i, \overline{P}_i) + Q^\mu P_\mu - f(\overline{Q}^\mu, P_\mu, \tau)\Big]_{\tau_1}^{\tau_2}, \end{aligned} \qquad (19)$$

where $\overline{Q}^i, \overline{P}_i, Q^\mu$ and P_μ must be written in terms of q^i and p_i. The action $\mathcal{S}[Q^i, P_i, N]$ describes an ordinary gauge system with the constraint $P_0 = 0$, so that the coordinate Q^0 is pure gauge, that is, Q^0 is not associated with a physical degree of freedom. This coordinate can be defined as an arbitrary function of τ by means of a canonical gauge choice. Writing Q^0 in terms of q^i and p_i, we have a function of the original phase space variables whose Poisson bracket with $H = P_0$ is positive definite; since H differs from the original Hamiltonian constraint only by a positive-definite function, we can always define a global phase time as

$$t(q^i, p_i) \equiv Q^0(q^i, p_i) \qquad (20)$$

because $[t(q^i, p_i), H(q^i, p_i)] = [Q^0, P_0] = 1$, and then

$$[t(q^i, p_i), \mathcal{H}(q^i, p_i)] > 0. \qquad (21)$$

The key point that allows us to define a global phase time for the minisuperspace by imposing a canonical gauge condition on the associated gauge system described by (Q^i, P_i) is that, in terms of these variables, we have a natural choice of a function whose Poisson bracket with the constraint is everywhere nonvanishing.

2.4. Path integral

The action $\mathcal{S}[Q^i, P_i, N]$ is stationary when the coordinates Q^i are fixed at the boundaries. The coordinates and momenta (Q^i, P_i) describe a gauge system with a linear constraint, so that this action allows one to obtain the amplitude of the transition $|Q_1^i, \tau_1\rangle \to |Q_2^i, \tau_2\rangle$ by the usual Faddeev–Popov procedure:

$$\langle Q_2^i, \tau_2 | Q_1^i, \tau_1 \rangle = \int DQ^0\, DP_0\, DQ^\mu\, DP_\mu\, DN \delta(\chi) |[\chi, P_0]| \mathrm{e}^{i\mathcal{S}[Q^i, P_i, N]} \qquad (22)$$

with $\mathcal{S}[Q^i, P_i, N]$ the gauge invariant action (18), and where $\chi = 0$ can be any canonical gauge condition. The Faddeev–Popov determinant $|[\chi, P_0]|$ ensures that the result does not depend on the gauge choice. If we perform functional integration on the lapse N enforcing the paths to lie on the constraint hypersurface $P_0 = 0$, we obtain

$$\begin{aligned} \langle Q_2^i, \tau_2 | Q_1^i, \tau_1 \rangle &= \int DQ^0\, DQ^\mu\, DP_\mu\, \delta(\chi)\, |[\chi, P_0]| \\ &\times \exp\left(i \int_{\tau_1}^{\tau_2} \left[P_\mu \frac{dQ^\mu}{d\tau} - h(Q^\mu, P_\mu, \tau) \right] d\tau \right), \end{aligned} \qquad (23)$$

where $h \equiv \partial f / \partial \tau$ is the true Hamiltonian of the reduced system. The path integral gives an amplitude between states characterized by the variables which, when fixed at the boundaries, make the action stationary. Since $\mathcal{S}$ is stationary when Q^i are fixed, we choose the gauge in the most general form giving Q^0 as a function of the other coordinates Q^μ and τ; thus a choice of the boundary values of the physical coordinates and τ fixes the boundary values of Q^0. With the choice $\chi \equiv Q^0 - T(Q^\mu, \tau) = 0$ and after trivially integrating in Q^0 we finally obtain

$$\begin{aligned} \langle Q_2^i, \tau_2 | Q_1^i, \tau_1 \rangle &= \int DQ^\mu\, DP_\mu \\ &\times \exp\left(i \int_{\tau_1}^{\tau_2} \left[P_\mu \frac{dQ^\mu}{d\tau} - h(Q^\mu, P_\mu, \tau) \right] d\tau \right), \end{aligned} \qquad (24)$$

so that we have $\langle Q_2^i, \tau_2 | Q_1^i, \tau_1 \rangle = \langle Q_2^\mu, \tau_2 | Q_1^\mu, \tau_1 \rangle$. Now, what we are in search for is an amplitude between states characterized by the original variables of the minisuperspace. Since the original action $S[q^i, p_i, N]$ is stationary when the coordinates q^i are fixed at the boundaries, it is common to seek a propagator of the form

$$\langle q_2^i | q_1^i \rangle, \tag{25}$$

so that the states are characterized only by the coordinates. But, as we have already remarked, in cosmology it is not always possible to define time in terms of the q^i only; then the amplitude $\langle q_2^i | q_1^i \rangle$ could not in general be understood as the probability that the observables of the system take a certain values at time t if at a previous time instant they took other given values.

If we pretend that

$$\langle Q_2^\mu, \tau_2 | Q_1^\mu, \tau_1 \rangle = \langle q_2^i | q_1^i \rangle,$$

the paths should be weighted by the action $\mathcal{S}$ in the same way as they are weighted by S, and the quantum states $|Q^\mu, \tau\rangle$ should be equivalent to $|q^i\rangle$. Since the path integral in the variables (Q^i, P_i) is gauge-invariant, this requirement is verified if it is possible to impose a *globally good* gauge condition $\tilde{\chi} = 0$ such that $\tau = \tau(q^i)$ is defined. But this can be fulfilled only when there exists a global time $t(q^i)$, which is not true in general. In the most general case, a global phase time must necessarily involve the momenta, and then we cannot fix the gauge in such a way that $\tau = \tau(q^i)$. Hence, we should admit the possibility of identifying the quantum states in the original phase space not by q^i but by a complete set of functions of both the coordinates and the momenta, q^i and p_i.

This may suggest giving up the idea of obtaining an amplitude for states characterized by the coordinates. However, while deparametrization in terms of the momenta may be completely valid at the classical level, at the quantum level there is an obstacle which is peculiar of gravitation [4]: There are basically two representations for quantum operators, the coordinate representation and the momentum representation, in which the states are characterized by occupation numbers associated with given values of the momenta. The latter is appropriate when the theory under consideration allows for the existence of asymptotically free states, so that there exists an interpretation in terms of creation and annihilation operators. In quantum cosmology such asymptotic states, in general, do not exist. An appropriate representation is then a coordinate one, in which the quantum states are represented by wave functions in terms of the coordinates. The usual Dirac–Wheeler–DeWitt quantization with momentum operators in the coordinate representation follows this line; but, as we have already noted, this formalism is devoid of a clear notion of time and evolution, unless there exists time in terms of the canonical coordinates only.

An intermediate way can then be followed: When the constraint allows for the existence of an intrinsic time, our deparametrization and path integral quantization procedure straightforwardly gives a transition amplitude for states characterized by the original coordinates; this provides a quantization with a clear distinction between time and the observables. On the other hand, when only an extrinsic time exists, we change from the original variables (q^i, p_i) to a set $(\tilde{q}^i, \tilde{p}_i)$ defined in such a way that the Hamiltonian constraint of a given model has a non-vanishing potential; then an intrinsic time exists in terms of the coordinates $\tilde{q}^i$, and the action $S[\tilde{q}^i, \tilde{p}_i, N]$ is stationary when $\tilde{q}^i$ are fixed at the boundaries. Therefore our procedure yields a transition amplitude for states characterized by the new coordinates, which is given by

$$\langle \tilde{q}_2^i | \tilde{q}_1^i \rangle = \langle Q_2^\mu, \tau_2 | Q_1^\mu, \tau_1 \rangle.$$

In both cases we obtain consistent quantization with a clear distinction between time and the observables. Though this seems to complicate the interpretation of the resulting propagator, the original momenta turn out to appear only in the time variable, while the new coordinates corresponding to the physical degrees of freedom depend on q^i only (a detailed discussion has been given in the context of quantization of the Taub anisotropic cosmology; see [22, 44]).

The form of the Hamiltonian h of the reduced system depends on the choice of the function f. We can choose f so that that the amplitude $\langle Q_2^\mu, \tau_2 | Q_1^\mu, \tau_1 \rangle$ is equivalent to $\langle \tilde{q}_2^i | \tilde{q}_1^i \rangle$. This requires that the Hamiltonian constraint should allow one to define time in terms of the coordinates $\tilde{q}^i$ and that the end point terms vanish on the constraint surface and in the gauge $\tilde{\chi} = 0$ defining $\tau = \tau(\tilde{q}^i)$, that is,

$$\left[\overline{Q}^i \overline{P}_i - W + Q^\mu P_\mu - f \right]_{\tau_1}^{\tau_2} \Bigg|_{P_0 = 0, \tilde{\chi} = 0} = 0. \tag{26}$$

Since the action $\mathcal{S}$ is gauge-invariant, this ensures that, with any gauge choice, the paths are weighted in the same way by $\mathcal{S}$ and S. This requirement gives a prescription for the generator $f(\overline{Q}^\mu, P_\mu, \tau)$ which determines the reduced Hamiltonian $h = \partial f / \partial \tau$. Since f depends only on the observables, h commutes with the complete Hamiltonian $K = NP_0 + h$, so that

$$\frac{dh}{d\tau} = \frac{\partial^2 f}{\partial \tau^2}.$$

Thus a generator f linear in τ yields a conserved Hamiltonian for the reduced system.

The reduced Hamiltonian h could be both positive or negative definite. As we shall illustrate with the second example in the next subsection, in general the sign of h will be in correspondence with the sign of a non-vanishing momentum of the set $\{\tilde{p}_i\}$ in terms of which the constraint surface splits into two sheets. The formalism will therefore include two theories for the physical degrees of freedom, corresponding to each sign of h associated with one of the two sheets of the constraint surface. The path integral in the reduced space will give two propagators, one for the evolution of the wave functions of each theory (see [4], and [24] for an analogous point of view). Note that then, if our path integral is to be associated with canonical quantization, splitting the formulation into two disjoint theories is in correspondence with two Schrödinger equations; so in general it does not coincide with the conventional Wheeler–DeWitt quantization. However, the existence of two disjoint theories, one for each sheet of the constraint surface, is a general property resulting from working with the time $t(\tilde{q}^i)$, which comes from the fact that we want to identify the path integral in the variables Q^i with a transition amplitude between states given in terms of the coordinates $\tilde{q}^i$; the non-equivalence between the Schrödinger and Wheeler–DeWitt quantizations, instead, depends on the model under consideration as well as on the choice of coordinates. This has been discussed in detail in [45].

2.5. Examples

Consider the Hamiltonian constraint of the most general empty homogeneous and isotropic cosmological model:

$$\mathcal{H} = -\frac{1}{4}\mathrm{e}^{-3\Omega}p_\Omega^2 - k\mathrm{e}^{\Omega} + \Lambda\mathrm{e}^{3\Omega} = 0. \tag{27}$$

This Hamiltonian corresponds to a universe with arbitrary curvature $k = -1, 0, 1$ and non-zero cosmological constant; we shall assume $\Lambda > 0$. If $k = 0$, we have a de Sitter universe. The classical evolution corresponds to an exponential expansion. For both $k = 0$ and $k = -1$, the potential is never zero, and thus p_Ω cannot change its sign. On the contrary, for the closed model $p_\Omega = 0$ is possible.

It is convenient to work with the rescaled Hamiltonian $H = \mathrm{e}^{-\Omega}\mathcal{H}$:

$$H = -\frac{1}{4}\mathrm{e}^{-4\Omega}p_\Omega^2 - k + \Lambda\mathrm{e}^{2\Omega} = 0. \tag{28}$$

The constraints H and $\mathcal{H}$ are equivalent because they differ only in a positive factor. The $\tau-$independent Hamilton–Jacobi equation for the Hamiltonian H has the solution

$$W(\Omega, \overline{P}_0) = 2\eta\int d\Omega\mathrm{e}^{2\Omega}\sqrt{\Lambda\mathrm{e}^{2\Omega} - k - \overline{P}_0}, \tag{29}$$

which is the generating function of the canonical transformation $(\Omega, \pi_\Omega) \to (\overline{Q}^0, \overline{P}_0)$ defined by

$$\overline{Q}^0 = -\eta\Lambda^{-1}\sqrt{\Lambda\mathrm{e}^{2\Omega} - k - \overline{P}_0}, \qquad \overline{P}_0 = H, \tag{30}$$

with $\eta = \mathrm{sign}(p_\Omega)$. Then we define the function $F = \overline{Q}^0 P_0 + f(\tau)$ which generates the second canonical transformation yielding a non-vanishing true Hamiltonian $h = \partial f/\partial\tau$ and $Q^0 = \overline{Q}^0$, $\overline{P}_0 = P_0$.

The variables Q^0 and P_0 describe the gauge system into which the model has been turned. The gauge can now be fixed by means of a $\tau-$dependent canonical condition like $\chi \equiv Q^0 - T(\tau) = 0$ with T a monotonic function of τ. Then we can define time as

$$t = Q^0|_{P_0=0} = -\eta\Lambda^{-1}\sqrt{\Lambda\mathrm{e}^{2\Omega} - k}, \tag{31}$$

or, using the constraint equation,

$$t(\Omega, p_\Omega) = -\frac{1}{2}\Lambda^{-1}\mathrm{e}^{-2\Omega}p_\Omega, \tag{32}$$

which is in agreement with the time obtained by matching the model with an ideal clock [7, 10]. There arises an important difference between the cases $k = -1$ and $k = 1$: for $k = -1$, the constraint surface splits into two disjoint sheets. In this case the evolution can be parametrized by a function of the coordinate Ω only, the choice given by the sheet on which the system remains: on the sheet $p_\Omega > 0$, the time is $t = -\Lambda^{-1}\sqrt{\Lambda\mathrm{e}^{2\Omega} + 1}$, while on the sheet $p_\Omega < 0$ we have $t = \Lambda^{-1}\sqrt{\Lambda\mathrm{e}^{2\Omega} + 1}$. Deparametrization of the flat model is completely analogous. For the closed model, instead, the potential can be zero, and the topology of the constraint surface is no more equivalent to that of two disjoint planes. Although for $\Omega = -\ln(\sqrt{\Lambda})$ we have $V(\Omega) = 0$ and $p_\Omega = 0$, at this point it is $dp_\Omega/d\tau \neq 0$. Hence in this case Ω cannot parametrize the evolution, because the system can pass from (Ω, p_Ω) to $(\Omega, -p_\Omega)$; therefore a global phase time must necessarily be defined as a function of both the coordinate and the momentum.

The system has one degree of freedom and one constraint, so that it is pure gauge. In other words, there is only one physical state: from a given point in phase space, any other point on the constraint surface can be reached by means of a finite gauge transformation. This provides a proof of the consistency of our procedure: it should be possible to verify that the transition probability written in terms of the variables which include a globally well defined time is equal to unity.

Quantization is straightforward, and the above observation is reflected in that we obtain the propagator [21]

$$\langle Q_2^0, \tau_2 | Q_1^0, \tau_1\rangle = \exp\left(-i\int_{\tau_1}^{\tau_2}\frac{\partial f}{\partial\tau}d\tau\right), \tag{33}$$

and then the transition probability from Q_1^0 at τ_1 to Q_2^0 at τ_2 is indeed

$$\left|\langle\, Q_2^0, \tau_2 | Q_1^0, \tau_1\,\rangle\right|^2 = 1. \tag{34}$$

When the model is open or flat, the coordinates Ω and Q^0 are uniquely related; hence the result simply reflects the fact that once a gauge is fixed, there is only one possible value of the scale factor $a \sim \mathrm{e}^{\Omega}$ at each τ, and

$$|\langle\Omega_2|\Omega_1\rangle|^2 = 1. \tag{35}$$

But in the case of a closed model, at each τ there are two possible values of the coordinate Ω; however, there is only one possible value of the momentum p_Ω at each τ. Hence the transition probability in terms of Q^0 does not correspond to the evolution of the coordinate Ω, but rather of its derivative, and the amplitude $\langle Q_2^0, \tau_2 | Q_1^0, \tau_1\rangle$ corresponds to the amplitude $\langle p_{\Omega,2} | p_{\Omega,1}\rangle$, and we have

$$|\langle p_{\Omega,2} | p_{\Omega,1}\rangle|^2 = 1. \tag{36}$$

The fact that the resulting amplitude is not equivalent to $\langle\Omega_2|\Omega_1\rangle$ is clearly not a failure of the quantization procedure, because for this model a characterization of states in terms of only the original coordinates is not possible if we want to retain a formally right notion of time for the whole evolution.

Now let us apply our formulation to a system with true degrees of freedom; consider a Hamiltonian constraint of the form

$$\mathcal{H} = G(\tilde{q}^2)(\tilde{p}_1^2 - \tilde{p}_2^2) + V(\tilde{q}^1, \tilde{q}^2) = 0, \tag{37}$$

where $G(\tilde{q}^2) > 0$. This constraint includes homogeneous and isotropic models, both relativistic and dilatonic, and also some anisotropic models, like Bianchi type I, the Kantowski–Sachs universe and the Taub universe (after an appropriate canonical transformation introduced in [22]). We shall restrict our analysis to the cases in which the potential $V(\tilde{q}^1, \tilde{q}^2)$ has a definite sign, so that $\tilde{q}^i$ is a set of coordinates including global time; we shall assume $V > 0$. We shall also suppose that the coordinates

$$x = x(\tilde{q}^1 + \tilde{q}^2), \qquad y = y(\tilde{q}^1 - \tilde{q}^2) \tag{38}$$

can be introduced, so that $4(\partial x/\partial \tilde{q}^1)(\partial y/\partial \tilde{q}^1) = V/G$; then we can write the constraint in the (scaled) equivalent form

$$H = p_x p_y + 1 = 0. \tag{39}$$

The solution of the corresponding Hamilton–Jacobi equation can be chosen in such a way that the canonical variables of the associated gauge system are given by

$$\begin{aligned} Q^0 &= y/P, \\ Q &= x + [y(1 - P_0) - \eta T(\tau)]/P^2, \\ P_0 &= p_x p_y + 1, \\ P &= p_x. \end{aligned} \tag{40}$$

Thus the canonical gauge condition $\chi \equiv Q^0 - T(\tau) = 0$ is associated with the extrinsic time $t = y/p_x$. We can also define an intrinsic time, which is related to the obtention of a transition amplitude between states characterized by the coordinates. The end point terms associated with the canonical transformation $(x, y, p_x, p_y) \to (Q^i, P_i)$ are of the form

$$B(\tau) = 2Q^0 - Q^0 P_0 - 2\eta \frac{T(\tau)}{P}. \tag{41}$$

On the constraint surface $P_0 = 0$, these terms clearly vanish in the gauge $\chi \equiv \eta Q^0 P - T(\tau) = 0$, which is in correspondence with the intrinsic time and the true Hamiltonian(s)

$$t(\tilde{q}^1, \tilde{q}^2) = \eta\, y(\tilde{q}^1 - \tilde{q}^2), \qquad h(Q, P, \tau) = \frac{\eta}{P}\frac{dT}{d\tau}, \tag{42}$$

with $\eta = \mathrm{sign}(p_x) = \mathrm{sign}(\tilde{p}_1 + \tilde{p}_2) = \mathrm{sign}(\tilde{p}_2)$, because $V > 0$ ensures that $|\tilde{p}_2| > |\tilde{p}_1|$. The propagator for the transition $|\tilde{q}_1^1, \tilde{q}_1^2\rangle \to |\tilde{q}_2^1, \tilde{q}_2^2\rangle$ is given by

$$\langle \tilde{q}_2^1, \tilde{q}_2^2 | \tilde{q}_1^1, \tilde{q}_1^2 \rangle = \int DQ\, DP \exp\left[i \int_{T_1}^{T_2} \left(P dQ - \frac{\eta}{P} dT \right) \right], \tag{43}$$

where the end points are given by $T_1 = \pm y(\tilde{q}_1^1 - \tilde{q}_1^2)$ and $T_2 = \pm y(\tilde{q}_2^1 - \tilde{q}_2^2)$. Note that with the gauge choice defining an intrinsic time, the observable Q reduces to a function of only the original coordinates:

$$Q|_{\chi=0} = x(\tilde{q}^1 + \tilde{q}^2).$$

Hence the paths go from $Q_1 = x(\tilde{q}_1^1 + \tilde{q}_1^2)$ to $Q_2 = x(\tilde{q}_2^1 + \tilde{q}_2^2)$. The propagator in the reduced space is therefore that of a system with a true degree of freedom given by the coordinate Q. Also, since $\tilde{p}_2$ does not vanish on the constraint surface, the coordinate $\tilde{q}^2$ is itself a global time, namely t^*; hence, though $\tilde{q}^2$ is not the time parameter in the path integral, the transition amplitude could be written as $\langle x_2, t_2^* | x_1, t_1^* \rangle$. Observe that by considering both possible signs of the reduced Hamiltonian, this path integral gives the transition amplitude for both theories corresponding to both sheets of the constraint surface identified by the sign of the momentum $\tilde{p}_2$.

3. A closed universe as a system without asymptotic states and the problem of time

Several years ago, a new approach to constructing quantum geometrodynamics was proposed by Savchenko, Shestakova and Vereshkov [32–36]. A central place in this approach is given to the Schrödinger equation for the wave function of the Universe which contains time as an external parameter, as in ordinary quantum mechanics. However, the appearance of time in the Schrödinger equation is a consequence of breaking down the gauge invariance of the theory. The proposed formulation is radically different from the generally accepted Wheeler–DeWitt quantum geometrodynamics, so one needs to have strong grounds for justifying this formulation.

3.1. Asymptotic states and gauge invariance

A key point of the authors' argument is an analysis of the role of asymptotic states in quantum gravity [34]. It is emphasized that any gauge-invariant quantum field theory is essentially based on the assumption about asymptotic states. Indeed, in the case of canonical quantization, in order to separate true physical degrees of freedom from "nonphysical" ones, we need to resolve gravitational constraints. It can be done in the limits of perturbation theory in asymptotically flat spaces or in some special cases. But in a general situation, if the Universe has some nontrivial topology and does not possess asymptotic states, this procedure meets unsurmountable mathematical difficulties.

In the path integral approach, which was accepted by the authors as the most adequate one, asymptotic boundary conditions ensure the BRST-invariance of a path integral and play the role of selection rules; as a consequence, the path integral does not depend on a gauge-fixing function (see [27]). Since a closed universe is a system without asymptotic states, it is not correct in this case to impose asymptotic boundary conditions in a path integral, so that the set of all possible transition amplitudes determined through the path integral inevitably involves gauge-noninvariant ones.

If the path integral is considered without asymptotic boundary conditions, it should be skeletonized on a *full* set of gauge-noninvariant equations obtained by varying an appropriate effective action including the ghost and gauge-fixing terms. Furthermore, there are two non-equivalent ways to proceed: to make use of the Batalin–Vilkovisky (Lagrangian) [6] or the Batalin–Fradkin–Vilkovisky (Hamiltonian) [18, 5, 19] effective action. There is a difference in the structure of ghost sectors, which, in turn, results from the fact that the gauge group of gravity does not coincide with the group of canonical transformations generated by the gravitational constraints. Two formulations based on the Lagrangian and Hamiltonian effective actions could be made equivalent in the gauge-invariant sector, the latter being singled out by means of asymptotic boundary conditions. Again, here we can see a crucial role of the assumption about asymptotic states in ensuring gauge invariance. In a situation without asymptotic states, one has to make a choice between these two effective actions; the authors give preference to the Lagrangian formalism since it maintains the original group of gauge transformations. Moreover, one cannot ensure the BRST invariance of the action without imposing asymptotic boundary condition, and the BFV scheme is then broken.

This approach leads to an extended set of the Einstein equations in which the constraints are broken even at the classical level. Eventually, this causes a dynamical Schrödinger equation and the appearance of time. A similar modification of the Hamiltonian constraint and the related time-dependent Schrödinger equation was discussed early by Weinberg [48] and Unruh [47]. The modification aimed at solving the cosmological constant problem, and the cosmological constant appeared to be a Hamiltonian eigenvalue. It resulted from an additional condition on the metric tensor which did not fix a gauge. It is then clear that the modification suggested by Weinberg and Unruh did not involve other gravitational constraints and equations of motion and was considered as a remedy for a particular (though very important) problem.

Another argument in favour of the gauge-noninvariant approach is the parametrization non-invariance of the Wheeler–DeWitt equation [26, 25]. The authors consider a unified interpretation of the choice of gauge variables (parametrization) and the choice of gauge conditions; the latter together determine the equations for the metric components $g_{0\mu}$, fixing a reference frame, as is illustrated by the scheme [37]

$$\begin{array}{ccc} \text{Parametrization} & & \\ g_{0\mu} = v_\mu\left(\mu_\nu, \gamma_{ij}\right) & & \\ + & \Rightarrow & \text{Equations for } g_{0\mu} \\ \text{Gauge conditions} & & g_{0\mu} = v_\mu\left(f_\nu\left(\gamma_{ij}\right), \gamma_{ij}\right) \\ \mu_\nu = f_\nu\left(\gamma_{ij}\right) & & \end{array}$$

Here μ_ν are new gauge variables, in particular, the lapse and shift functions, N and N_i, γ_{ij} is the 3-metric. Thus even if one considers μ_ν to be independent of γ_{ij}, different parametrizations will correspond to different reference frames. This leads to the conclusion that a transition to another gauge variable is formally equivalent to imposing a new gauge condition, and vice versa, and the parametrization non-invariance of the Wheeler–DeWitt equation is ill-hidden gauge non-invariance.

3.2. Hamiltonian dynamics in extended phase space

After these preliminary notes, let us go into mathematical details. The authors consider a simple minisuperspace model with the gauged action

$$S = \int dt \left\{ \frac{1}{2} v(\mu, Q^a) \gamma_{ab} \dot{Q}^a \dot{Q}^b - \frac{1}{v(\mu, Q^a)} U(Q^a) + \pi \left(\dot{\mu} - f_{,a} \dot{Q}^a \right) - i w(\mu, Q^a) \dot{\bar{\theta}} \dot{\theta} \right\}. \tag{44}$$

Here Q^a stands for physical variables such as a scale factor or gravitational-wave degrees of freedom and material fields, and an arbitrary parametrization of a gauge variable μ determined by the function $v(\mu, Q^a)$ is accepted. For example, in the case of an isotropic universe or the Bianchi IX model μ is bound to the scale factor r and the lapse function N by the relation

$$\frac{r^3}{N} = v(\mu, Q^a). \tag{45}$$

A special class of time-independent gauges is used:

$$\mu = f(Q^a) + k; \qquad k = \text{const.} \tag{46}$$

It is convenient to present the gauge in a differential form,

$$\dot{\mu} = f_{,a} \dot{Q}^a, \qquad f_{,a} \stackrel{\text{def}}{=} \frac{\partial f}{\partial Q^a}. \tag{47}$$

Here θ, $\bar{\theta}$ are the Faddeev–Popov ghosts after the replacement $\bar{\theta} \to -i\bar{\theta}$. Furthermore,

$$w(\mu, Q^a) = \frac{v(\mu, Q^a)}{v_{,\mu}}; \qquad v_{,\mu} \stackrel{\text{def}}{=} \frac{\partial v}{\partial \mu}. \tag{48}$$

Varying the effective action (44) with respect to Q^a, μ, π and θ, $\bar{\theta}$ one gets, accordingly, the equations of motion for physical variables, the constraint, the gauge condition and equations for ghosts. The extended set of Lagrangian equations is complete in the sense that it enables one to formulate the Cauchy problem. An explicit substitution of trivial solutions for ghosts and the Lagrangian multiplier π to this set of equations turns one back to the gauge-invariant classical Einstein equations.

The path integral approach does not require the construction of a Hamiltonian formulation before deriving the Schrödinger equation, but it implies that the Hamiltonian formulation can be constructed. Indeed, in the class of gauges (47) the Hamiltonian can be obtained in a usual way, according to the rule $H = p\dot{q} - L$, where (p, q) are the canonical pairs of extended phase space (EPS), by introducing momenta conjugate to all degrees of freedom including the gauge ones,

$$H = P_a\dot{Q}^a + \pi\dot{\mu} + \bar{\rho}\dot{\theta} + \dot{\bar{\theta}}\rho - L$$

$$= \frac{1}{2}G^{\alpha\beta}P_\alpha P_\beta + \frac{1}{v(\mu, Q^a)}U(Q^a) - \frac{i}{w(\mu, Q^a)}\bar{\rho}\rho, \quad (49)$$

where $\alpha = (0, a),\ Q^0 = \mu$,

$$G^{\alpha\beta} = \frac{1}{v(\mu, Q^a)}\begin{pmatrix} f_{,a}f^{,a} & f^{,a} \\ f^{,a} & \gamma^{ab} \end{pmatrix}. \quad (50)$$

The set of Hamiltonian equations in EPS

$$\dot{X} = \{H, X\}, \quad X = (P_a,\ Q^a,\ \pi,\ \mu,\ \bar{\rho},\ \theta,\ \rho,\ \bar{\theta}) \quad (51)$$

is completely equivalent to the extended set of Lagrangian equations, the constraint and the gauge condition acquiring the status of Hamiltonian equations. The idea of extended phase space is exploited in the sense that the gauge and ghost degrees of freedom are treated on an equal basis with other variables. This gave rise to the name “quantum geometrodynamics in extended phase space” accepted by the authors.

Obviously, Hamiltonian dynamics is constructed here in a different way than in the BFV approach where constraints are maintained and play a central role. Here the Hamiltonian constraint is modified and looks as follows:

$$\dot{\pi} = -\frac{i}{w^2(\mu, Q^a)}w_{,\mu}\bar{\rho}\rho + \frac{1}{v^2(\mu, Q^a)}v_{,\mu} \times$$
$$\times \left[\frac{1}{2}\left(P_aP^a + 2\pi f_{,a}P^a + \pi^2 f_{,a}f^{,a}\right) + U(Q^a)\right]. \quad (52)$$

The gauge-dependent terms can be eliminated by making use of trivial solutions for π and the ghosts. It is generally accepted that these trivial solutions can be singled out by asymptotic boundary conditions, if one ignores the problem of Gribov's copies. Furthermore, the restriction on the class of admissible parametrizations, $v(\mu, Q) = \frac{u(Q)}{\mu}$, together with the trivial solutions for π and ghosts reduce the constraint to the form

$$\mathcal{T} = \frac{1}{2u(Q^a)}P_aP^a + \frac{1}{u(Q^a)}U(Q^a) = 0. \quad (53)$$

The restriction on the class of parametrizations is necessary for the physical Hamiltonian to be proportional to the constraint $H_0 = \mu\mathcal{T}$. We come to the conclusion that the primary and secondary constraints $\pi = 0$, $\mathcal{T} = 0$ correspond to a particular situation when it is possible to single out the trivial solutions for π and ghosts by asymptotic boundary conditions. So, in this sense, the Dirac quantization is applicable to systems with asymptotic states only. The same is true for the BFV approach which inherits most of features of the Dirac quantization.

3.3. Quantum geometrodynamics in extended phase space

The constraint (52) can be presented in the form $H = E$, where E is a conserved quantity (a new integral of motion). Accordingly, in quantum theory the relation $H = E$ should be replaced by a stationary Schrödinger equation, $H|\Psi\rangle = E|\Psi\rangle$, the Hamiltonian spectrum in EPS being not limited by the unique zero eigenvalue.

So, there is no reason to require that a wave function of a closed universe should satisfy the Wheeler–DeWitt equation. Independently of our notion of gauge invariance or non-invariance of the theory, the wave function should obey some Schrödinger equation. The Schrödinger equation is derived from the path integral with the effective action (44) by a standard method originated by Feynman [16, 9]. For the present model it reads

$$i\,\frac{\partial\Psi(\mu, Q^a, \theta, \bar{\theta};\, t)}{\partial t} = H\Psi(\mu,\ Q^a,\ \theta,\ \bar{\theta};\ t), \quad (54)$$

where

$$H = -\frac{i}{w}\frac{\partial}{\partial\theta}\frac{\partial}{\partial\bar{\theta}} - \frac{1}{2M}\frac{\partial}{\partial Q^\alpha}MG^{\alpha\beta}\frac{\partial}{\partial Q^\beta} + \frac{1}{v}(U - V); \quad (55)$$

the operator H corresponds to the Hamiltonian in EPS (49). M is the measure in the path integral,

$$M(\mu, Q^a) = v^{\frac{K}{2}}(\mu, Q^a)w^{-1}(\mu, Q^a); \quad (56)$$

K is the number of physical degrees of freedom; the wave function is defined on extended configurational space with the coordinates $\mu,\ Q^a,\ \theta,\ \bar{\theta}$. V is a quantum correction to the potential U, which depends on the chosen parametrization (45) and gauge (46):

$$V = \frac{5}{12w^2}\left(w_{,\mu}^2 f_{,a}f^{,a} + 2w_{,\mu}f_{,a}w^{,a} + w_{,a}w^{,a}\right)$$
$$+ \frac{1}{3w}\left(w_{,\mu,\mu}f_{,a}f^{,a} + 2w_{,\mu,a}f^{,a} + w_{,\mu}f^{;a}_{,a} + w^{;a}_{,a}\right)$$
$$+ \frac{K-2}{6vw}(v_{,\mu}w_{,\mu}f_{,a}f^{,a} + v_{,\mu}f_{,a}w^{,a}$$
$$+ w_{,\mu}f_{,a}v^{,a} + v_{,a}w^{,a})$$
$$- \frac{K^2 - 7K + 6}{24v^2}\left(v_{,\mu}^2 f_{,a}f^{,a} + 2v_{,\mu}f_{,a}v^{,a} + v_{,a}v^{,a}\right)$$
$$+ \frac{1-K}{6v}(v_{,\mu,\mu}f_{,a}f^{,a} + 2v_{,\mu,a}f^{,a} + v_{,\mu}f^{;a}_{,a} + v^{;a}_{,a}). \quad (57)$$

Let us emphasize that the Schrödinger equation (54)–(57) is *a direct mathematical consequence* of a path integral with the effective action (44) without asymptotic boundary conditions. Once we reject imposing asymptotic boundary conditions, we *are doomed* to come to a gauge-dependent description of the Universe.

The general solution to the Schrödinger equation has the following structure:

$$\Psi(\mu,\, Q^a,\, \theta,\, \bar{\theta};\, t) = \int \Psi_k(Q^a,\, t)\, \delta(\mu - f(Q^a) - k)\, (\bar{\theta} + i\theta)\, dk. \qquad (58)$$

As one can see, the general solution is a superposition of eigenstates of a gauge operator,

$$\{\mu - f(Q^a)\}|k\rangle = k\,|k\rangle; \quad |k\rangle = \delta\,(\mu - f(Q^a) - k)\,. \qquad (59)$$

It can be interpreted in the spirit of Everett's "relative state" formulation. In fact, each element of the superposition (58) describes a state in which the only gauge degree of freedom μ is definite, so that a time scale is determined by processes in the physical subsystem through the functions $v(\mu,\, Q^a)$, $f(Q^a)$ (see (45), (46)), while k is determined by the initial clock setting. The function $\Psi_k(Q^a,\, t)$ describes a state of the physical subsystem for a reference frame fixed by the condition (46). It is a solution to the equation

$$i\,\frac{\partial \Psi_k(Q^a;\, t)}{\partial t} = H_{(\mathrm{phys})}[f]\Psi_k(Q^a;\, t), \qquad (60)$$

$$H_{(\mathrm{phys})}[f] = \left[-\frac{1}{2M}\frac{\partial}{\partial Q^a}\frac{1}{v}M\gamma^{ab}\frac{\partial}{\partial Q^b} + \frac{1}{v}(U - V)\right]\Bigg|_{\mu = f(Q^a) + k}. \qquad (61)$$

One can seek a solution to Eq. (60) in the form of a superposition of stationary state eigenfunctions:

$$\Psi_k(Q^a,\, t) = \sum_n c_{kn}\psi_n(Q^a)\exp(-iE_n t);$$
$$H_{(\mathrm{phys})}[f]\psi_n(Q^a) = E_n\psi_n(Q^a). \qquad (62)$$

The eigenvalues E_n should not be associated with the energy of any material field. It results from fixing a gauge condition and characterizes a subsystem which corresponds to observational means — a reference frame (see [34, 35] for details).

Having constructed the general solution to the Schrödinger equation, one can pose the question: can a physical part of the wave function obey the Wheeler–DeWitt equation under some additional conditions? A natural additional condition in EPS is the requirement of BRST invariance of the wave function. Indeed, in the BFV approach the requirement of BRST invariance leads immediately to the Wheeler–DeWitt equation. The BRST charge has an especially simple form for the present model,

$$\Omega_{BFV} = \eta^{\alpha}\mathcal{G}_{\alpha} = \mathcal{T}\theta - i\pi\rho, \qquad (63)$$

where $\mathcal{G}_{\alpha} = (\pi,\, \mathcal{T})$ is the full set of constraints, and due to arbitrariness of BFV ghosts $\{\eta^{\alpha}\}$ one gets the Wheeler – DeWitt equation $\mathcal{T}\,|\Psi\rangle = 0$ from the requirement $\Omega_{BFV}\,|\Psi\rangle = 0$.

It is not the case in the approach considered above. We should recall that the original group of transformations was the group of gauge transformations in the Lagrangian formalism. It is the reason why the transformations generated by (63) do not coincide with those under which the action (44) is invariant. The BRST charge constructed according to the BFV prescription turns out to be irrelevant in this consideration. Instead, there exists another quantity that plays the role of the BRST generator,

$$\Omega = w(Q^a, \mu)\;\pi\dot{\theta} - H\theta = -\,i, \pi\rho - H\theta. \qquad (64)$$

It is easy to check that (64) generates transformations in EPS which are identical to the BRST transformations in the Lagrangian formalism. Nevertheless, it cannot be presented as a combination of constraints with infinitesimal parameters replaced by ghosts and cannot help us to obtain the Wheeler–DeWitt equation [35].

On the other hand, as the authors show, the fact that the wave function obeys the Wheeler–DeWitt equation does not mean that this wave function describes the Universe in a gauge-invariant way, i.e., independently of a reference frame. If one puts $\mu = k$, $E = 0$ and restricts the class of parametrizations, as was done above (see (53) and the text before), the equation for the physical part of the wave function $H_{(\mathrm{phys})}\Psi_k(Q^a) = E\Psi_k(Q^a)$ is reduced to the Wheeler–DeWitt equation with its parametrization non-invariance and without any visible vestige of a gauge. By construction, however, a solution to this equation corresponds to a particular choice of a gauge condition and a particular line in the Hamiltonian spectrum. It is enough then to fix the parametrization to complete the choice of a reference frame. It confirms the conclusion about ill-hidden gauge-noninvariance of the Wheeler–DeWitt equation which has been done in the beginning of this section.

All this demonstrates that this attempt to derive a gauge-invariant quantum theory from a more general gauge non-invariant one rises many questions. For a system with asymptotic states, we have the BFV approach where we consider constraints at the classical level before quantization. But even in this case making use of asymptotic boundary conditions to exclude the gauge-non-invariant terms is an idealization in the sense that we neglect the problem of Gribov's copies.

3.4. Topology of the Universe and irreversibility of time

In conclusion we shall touch upon one of the consequences of the approach presented — *irreversibility* of a transition to another reference frame in the framework of the gauge-non-invariant description [38]. Since the reference frame was declared to be a constituent of an integrated system as well as the physical Universe and plays a role of a measuring device, any change of the reference frame will cause changes in the observed physical picture. Indeed, let us consider a small variation of the gauge-fixing function $f(Q^a)$, so that the reference frame will be fixed by the condition

$$\mu = f(Q^a) + \delta f(Q^a) + k. \tag{65}$$

Then, in a new basis corresponding to this reference frame, the wave function will take the form

$$\Psi(\mu,\, Q^a,\, \theta,\, \bar\theta;\, t) = \int \tilde\Psi_k(Q^a,\, t)\, \delta\big(\mu - f(Q^a) - \delta f(Q^a) - k\big)(\bar\theta + i\theta)\, dk. \tag{66}$$

Here the function $\tilde\Psi_k(Q^a,\, t)$ satisfies Eq. (60) with the Hamiltonian

$$H_{(\rm phys)}[f + \delta f] = \left[-\frac{1}{2M}\frac{\partial}{\partial Q^a}\left(\frac{1}{v}M\gamma^{ab}\frac{\partial}{\partial Q^b}\right) + \frac{1}{v}(U - V)\right]\Bigg|_{\mu = f(Q^a)+\delta f(Q^a)+k}. \tag{67}$$

Obviously, the equation for the physical part of the wave function with the Hamiltonian (67) cannot be reduced, in general, to an equation with the Hamiltonian (61). The measure in the subspace of physical degrees of freedom also depends on the gauge condition chosen, as follows from the normalization equation

$$\int \Psi^*_{k'}(Q^a,\, t)\, \Psi_k(Q^a,\, t)\, \delta(\mu - f(Q^a) - k') \times \delta(\mu - f(Q^a) - k)\, dk'\, dk\, M(\mu,\, Q^a)\, d\mu \prod_a dQ^a = \int \Psi^*_k(Q^a,\, t)\, \Psi_k(Q^a,\, t)\, M(f(Q^a) + k,\, Q^a) \times \prod_a dQ^a\, dk = 1. \tag{68}$$

Due to smallness of $\delta f(Q^a)$, one can write

$$H_{(\rm phys)}[f + \delta f] = H_{(phys)}[f] + W[\delta f] + V_1[\delta f]. \tag{69}$$

For our minisuperspace model, the operator $W[\delta f]$ reads

$$W[\delta f] = \left\{\frac{1}{2M^2}\frac{\partial M}{\partial \mu}\delta f \frac{\partial}{\partial Q^a}\left(\frac{1}{v}M\gamma^{ab}\frac{\partial}{\partial Q^b}\right) - \frac{1}{2M}\frac{\partial}{\partial Q^a}\left[\left(\frac{1}{v}\frac{\partial M}{\partial \mu} - \frac{M}{v^2}\frac{\partial v}{\partial \mu}\right) \times \delta f \gamma^{ab}\frac{\partial}{\partial Q^b}\right]\right\}\Bigg|_{\mu = f(Q^a)+k}, \tag{70}$$

and $V_1[\delta f]$ is the change of the quantum potential V (57) in the first order of δf.

One can inquire how the probabilities of stationary states (62) change under the perturbation $W[\delta f] + V_1[\delta f]$, which is due to a small variation of the gauge-fixing function $f(Q^a)$. The Hamiltonian (67) is Hermitian by construction in a space with the measure $M(f(Q^a) + \delta f(Q^a) + k,\, Q^a)$, however, it is not Hermitian in a space with the measure $M(f(Q^a) + k,\, Q^a)$ in which the functions (62) are normalized. In this space the operator (70) will have, in general, an anti-Hermitian part. So any transition to another reference frame must be *irreversible.*

This is true for a transition to another reference frame in the same spacetime region, and this is also true if spacetime consists of several regions where different reference frames are introduced. A nontrivial topology of the Universe may be a reason why one has to introduce various reference frames in different spacetime regions. In particular, we can consider mutually intersecting spacetime regions ordered in time. Every time when we move from one region to another, the physical part of the wave function will undergo a non-unitary transformation followed by changing in the measure in the subspace of physical degrees of freedom, which may lead to irreversible consequences in the physical picture of the Universe. If so, taking into account an interaction with the reference frame — the measuring instrument representing the observer in quantum theory of gravity — not only enables one to introduce time into quantum geometrodynamics, but also may attach an irreversible nature to the cosmological evolution.

4. Discussion

Among many attempts to give a solution to the problem of time, we have paid a considerable attention to two approaches, which are, as a matter of fact, very different. The first one, by Simeone and collaborators, considered in Sec. 2, is a development of the unitary approach to quantum gravity inspired by earlier works of Barvinsky and Hájíček. The key point here is a reduction of the gravitational action to that of an ordinary gauge system. Since the Hamiltonian constraint is quadratic in momenta, we come, in general, to two formulations of the theory which correspond to two disjoint sheets of the constraint surface given by the two signs of the momentum conjugated to a time variable. The proposed procedure enables one to formulate the theory in terms of true degrees of freedom and then to return to a transition amplitude between states characterized by the original variables of phase space, so that the whole scheme is gauge-invariant. This approach demonstrates how time can be introduced into the theory *without breaking down its gauge invariance*. Let us emphasize that the requirement of gauge invariance is conventionally thought of to be one of basic requirements for a physical theory.

In this sense, the second approach, presented in Sec. 3, is very radical. According to the analysis by Isham [30], approaches to the problem of time can be subdivided into three main categories: those in which time is identified before quantizing, those in which time is identified after quantizing and approaches in which time plays no fundamental role at all. The proposal by Savchenko, Shestakova and Vereshkov does not belong to any of these categories. In their scheme, time naturally appears while quantizing a gravitational system, namely, while driving a Schrödinger equation from the path integral. In this consideration, time has a status of an external parameter, as in ordinary quantum mechanics. The price for it is a refusal from gauge invariance of the theory.

We would note that the fact that the Universe does not possess asymptotic states has not been analysed earlier from the viewpoint of its connection with gauge invariance. Traditionally, the Universe was quantized as any gauge system dealt with in laboratory physics. On the other hand, in modern field theory one can find indications that the role of gauge degrees of freedom may not be just auxiliary. It will suffice to mention the Aharonov–Bohm effect and instanton solutions. All of them originate from a nontrivial topological structure of spacetime. A future development will give an objective appraisal to the proposed approaches to the problem of time which remains to be a fundamental problem in the construction of quantum gravity.

References

[1] A.O. Barvinsky and V.N. Ponomariov, *Phys. Lett.* **B 167**, 289 (1986).

[2] A.O. Barvinsky, *Phys. Lett.* **B 175**, 401 (1986).

[3] A.O. Barvinsky, *Phys. Lett.* **B 195**, 289 (1987).

[4] A.O. Barvinsky, *Phys. Rep.* **230**, 237 (1993).

[5] I.A. Batalin and G.A. Vilkovisky, *Phys. Lett.* **B 69**, 309 (1977).

[6] I.A. Batalin and G.A. Vilkovisky, *Phys. Lett.* **B 102**, 27 (1981).

[7] S.C. Beluardi and R. Ferraro, *Phys. Rev.* **D 52**, 1963 (1995).

[8] J.D. Brown and K.V. Kuchař, *Phys. Rev.* **D 51**, 5600 (1995)

[9] K.S. Cheng, *J. Math. Phys.* **13**, 1723 (1972).

[10] H. De Cicco and C. Simeone, *Gen. Rel. Grav.* **31**, 1225 (1999).

[11] H. De Cicco and C. Simeone, *Int. J. Mod. Phys.* **A 14**, 5105 (1999).

[12] P.A.M. Dirac, *Physik. Zeits. Sowjetunion* **3**, 64 (1933).

[13] L.D. Faddeev and V.N. Popov, *Phys. Lett.* **B 25**, 29 (1967).

[14] L.D. Faddeev and A.A. Slavnov, "Gauge Fields: Introduction to Quantum Theory", Benjamin/Cummings Publishing, 1980.

[15] R. Ferraro and C. Simeone, *J. Math. Phys,* **38**, 599 (1997).

[16] R.F. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).

[17] R.F. Feynman and A.R. Hibbs, "Quantum Mechanics and Path Integrals", McGraw–Hill, New York, 1965.

[18] E.S. Fradkin and G.A. Vilkovisky, *Phys. Lett.* **B 55**, 224 (1975).

[19] E.S. Fradkin and T.E. Fradkina, *Phys. Lett.* **B 72**, 343 (1978).

[20] G. Giribet and C. Simeone, *Mod. Phys. Lett.* **A 16**, 19 (2001).

[21] G. Giribet and C. Simeone, *Phys. Lett.* **A 287**, 344 (2001).

[22] G. Giribet and C. Simeone, *Int. J. Mod. Phys.* **A 17**, 2885 (2002).

[23] V.N. Gribov, *Nucl. Phys.* **B 139**, 1 (1978).

[24] P. Hájíček, *Phys. Rev.* **D 34**, 1040 (1986).

[25] J.J. Halliwell, *Phys. Rev.* **D 38**, 2468 (1988).

[26] S.W. Hawking and D.N. Page, *Nucl. Phys.* **B 264**, 185 (1986).

[27] M. Henneaux, *Phys. Rep.* **126**, 1 (1985).

[28] M. Henneaux, C. Teitelboim and J.D. Vergara, *Nucl. Phys.* **B 387**, 391 (1992).

[29] M. Henneaux and C. Teitelboim, "Quantization of Gauge Systems", Princeton University Press, New Jersey, 1992.

[30] C. Isham, "Canonical quantum gravity and the problem of time", lectures presented at NATO Advanced Study Institute, Salamanca, June 1992.

[31] L.D. Landau and E.M. Lifshitz, "Mechanics", Pergamon Press, Oxford, 1960.

[32] V.A. Savchenko, T.P. Shestakova and G.M. Vereshkov, *Int. J. Mod. Phys.* **A 14**, 4473 (1999).

[33] V.A. Savchenko, T.P. Shestakova and G.M. Vereshkov, *Int. J. Mod. Phys.* **A 15**, 3207 (2000).

[34] V.A. Savchenko, T.P. Shestakova and G.M. Vereshkov, *Grav. & Cosmol.* **7**, 18 (2001).

[35] V.A. Savchenko, T.P. Shestakova and G.M. Vereshkov, *Grav. & Cosmol.* **7**, 102 (2001).

[36] T.P. Shestakova, *Grav. & Cosmol.* **5**, 297 (1999).

[37] T.P. Shestakova, *in:* Proceedings of the IV International Conference "Cosmion-99", *Grav. & Cosmol.* **6**, Supplement, 47 (2000).

[38] T.P. Shestakova, *in:* Proceedings of the V International Conference on Gravitation and Astrophysics of Asian-Pacific countries, *Grav. & Cosmol.* **8**, Supplement II, 140 (2002).

[39] T.P. Shestakova and C. Simeone, "The Problem of Time and Gauge Invariance in the Quantization of Cosmological Models. I. Canonical Quantization Methods", *Grav. & Cosmol.* **10**, 161 (2004).

[40] C. Simeone, *J. Math. Phys.* **39**, 3131 (1998).

[41] C. Simeone, *J. Math. Phys.* **40**, 4527 (1999).

[42] C. Simeone, *Gen. Rel. Grav.* **32**, 1835 (2000).

[43] C. Simeone, *Gen. Rel. Grav.* **34**, 1887 (2002).

[44] C. Simeone, "Deparametrization and Path Integral Quantization of Cosmological Models" (World Scientific Lecture Notes in Physics 69), World Scientific, Singapore, 2002.

[45] C. Simeone, *Phys. Lett.* **A 310**, 143 (2003).

[46] C. Teitelboim, *Phys. Rev.* **D 25**, 3159 (1982).

[47] W.G. Unruh, *Phys. Rev.* **D 40**, 1048 (1989).

[48] S. Weinberg., *Rev. Mod. Phys.* **61**, 1 (1989).