

Quantum Cosmology

Nelson Pinto Neto

Centro Brasileiro de Pesquisas Físicas - CBPF/CNPq/Lafex
Rua Dr. Xavier Sigaud, 150 - Urca
22290-180, Rio de Janeiro, RJ, Brazil

ABSTRACT

The canonical quantization of General Relativity is summarized. Its application to quantum cosmology is described. The interpretation and time issues, and the problem of initial conditions are discussed. Some possible solutions to these problems are proposed.

Key-words: Quantum cosmology; Quantum gravity; Interpretation of quantum mechanics.

1 Introduction

The singularity theorems [1] show that, under reasonable physical assumptions, the universe has developed an initial singularity, which is called the big bang, and will develop future singularities in the form of black holes and, perhaps, of a big crunch. Until now, singularities are out of the scope of any physical theory. If we take the pretentious attitude that a physical theory can describe the whole universe at every instant, even at its moment of creation if it has had one, (which is the best attitude because it is the only way to seek the limits of physical science), we must assume that the ‘reasonable physical assumptions’ of the theorems are not valid under extreme situations of very high energy density and curvature, which is very plausible. We may say that general relativity or any other matter field theory must be changed under these extreme conditions. One good point of view (which is not the only one) is to think that quantum gravitational effects become important. We should then construct a quantum theory of gravitation and apply it to cosmology. It is a good point of view because, besides the possibility of obtaining from quantum gravity a solution to the singularity problem, we gain from quantum theory the possibility of constructing a theory of initial conditions for the universe. This theory could then explain why the universe is remarkably homogeneous and isotropic and even why the constants of nature have the values we observe they have. Moreover, it could give the spectrum of quantum fluctuations of geometry and matter of primordial origin and provide us with a complete theory of galaxy formation.

We call quantum cosmology [2, 3, 4] as this attempt to apply quantum gravity ideas to the universe as a whole. As we have seen, the goals of quantum cosmology are rather ambitious. However, its problems are pairwise with its ambition. This is because it tries to put together three of the major and revolutionary achievements of physical science in the twentieth century. One, general relativity, which describes gravitation, has abolished the concept of absolute spacetime by treating its metric as a dynamical variable in interaction with matter and with itself. Other, quantum mechanics, which is the correct description of matter at atomic scales or below, has put serious objections to the existence of a very natural and basic concept: objective reality. Finally, cosmology, is a theory of a single system, the universe as a whole, including us, observers, a situation which is very unfamiliar to natural science. That is why so much time was needed for scientists to insert cosmology in the domain of natural science. Hence, needless to say how hard it is to put these three theories together.

The most exciting feature of this union is that all the difficulties of particular theories that once were forgotten because they were not important for all practical purposes, become crucial in quantum cosmology. Let us begin with the marriage of quantum mechanics with general relativity (or some gravity theory which contains it). No quantum gravity theory was proved to be renormalizable. One good candidate is string theory, which has also the ambition of being a theory of everything. However, even if one of such theories is proven to be renormalizable, it must also be shown that the perturbation series of the theory does not diverge when summed over. In other quantum field theories, like in quantum electrodynamics, it is argued that such divergence involves very high energies that cannot be probed now and at Planckian scales some more fundamental theory must be used. This reasoning cannot be applied to quantum gravity; after all, quantum gravity is *the*

theory to be applied at Planckian scales. Nonperturbative quantum gravity has also a lot of unresolved problems (complicate constraint equations, lack of unitary evolution, etc).

The application of quantum gravity to cosmology adds new problems. How can we apply the standard Copenhagen interpretation to a single system? What happens with its probabilistic interpretation? Who are the observers of the whole universe? Where in a quantum universe can we find a classical domain where we could construct our classical apparatus, and test the theory? This is not a problem of quantum gravity alone because there is no problem with the concept of an ensemble of black holes and a classical domain outside it. Finally, in quantum mechanics, time, in spite of seeming to be a measurable physical quantity, is not treated as an observable (hermitean operator) but as an external evolution parameter (c-number). In the quantum cosmology of a closed universe, there is no place for an external parameter. So, what happens with time; does it become an operator? These are some of the difficult issues which the subject of quantum cosmology has to give an answer in order to have a meaning.

In these lectures we will try to explain some ideas on how can quantum cosmology achieve its ambitious goals and what are the attempts to answer some of the profound and difficult issues it has raised.

In the following section we will set the problem of the initial cosmological singularity and motivate the study of quantum cosmology. After, we will show that the Copenhagen interpretation of quantum mechanics cannot be used in quantum cosmology and we will present some of the alternative interpretations that can be consistent with a theory of the whole universe.

In the third section the canonical quantization of general relativity will be developed and the Wheeler-DeWitt equation obtained. The issue of time will be discussed and we will advocate the idea that time has no meaning at Planckian scales. It can be recovered only at the semi-classical limit.

In the fourth section we will begin to present quantum cosmology as a theory of initial conditions of the universe by using a modified version of the many-worlds interpretation of quantum mechanics. The idea is to find peaks of the semi-classical Wigner function in order to find the most probable cosmological classical solution. After introducing the notion of minisuperspace, we will present a particular example where this idea can be applied. Unfortunately, it will be shown that this program can not be implemented in general. Also, from the study of the minisuperspace model introduced in this section, it will be evident the need of having boundary conditions to the Wheeler-DeWitt equation, which yields the motivation for the next section.

In section 5, the no-boundary boundary condition to the Wheeler-DeWitt equation will be presented. It will be applied to the example of the preceding section and to other models. Here, quantum cosmology as a theory of initial conditions exhibits all its potentiality. In particular, we will indicate how the conditions of having inflation might be obtained, and how the spectrum of quantum mechanical perturbations responsible for galaxy formation might be attained. Of course these are not closed answers and we will show their problems. After, we will present the notion of decoherence and show how it can be useful in quantum cosmology. In particular, it may save the program of finding the most probable classical solution from the Wigner function and explain how a classical universe can emerge from a quantum one.

In section 6, we will introduce some other alternative interpretations of quantum mechanics: the formulation of quantum mechanics in terms of histories elaborated by Griffiths, Omnès, Hartle and Gell-Mann, and the ontological interpretation of Bohm-de Broglie-Healey, and their relevance to quantum cosmology.

Finally, in the last section, we will conclude with a summary of the stimulating results of quantum cosmology, the many problems that are still unresolved, and a personal point of view about the good directions that should be followed.

Conventions and notation

- a) metric signature: $(-, +, +, +)$
- b) greek indices vary from 0 to 3 and latin indices from 1 to 3
- c) four-dimensional covariant derivative of a four-vector A^β :

$$\nabla_\alpha A^\beta \equiv \partial_\alpha A^\beta + \Gamma_{\alpha\nu}^\beta A^\nu \quad (1)$$

where

$$\Gamma_{\alpha\nu}^\beta \equiv \frac{1}{2} g^{\beta\mu} (\partial_\nu g_{\alpha\mu} + \partial_\alpha g_{\nu\mu} - \partial_\mu g_{\alpha\nu}) \quad (2)$$

and $\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}$

- d) three-dimensional covariant derivative of a three-vector A^j :

$$D_i A^j \equiv \partial_i A^j + {}^3\Gamma_{ik}^j A^k \quad (3)$$

where

$${}^3\Gamma_{ik}^j = \frac{1}{2} h^{jl} (\partial_i h_{kl} + \partial_k h_{il} - \partial_l h_{ik}) \quad (4)$$

where h_{ij} is a 3-dimensional metric and h^{ij} its inverse.

- e) four-dimensional curvature:

$$R_{\nu\alpha\beta}^\mu \equiv \nabla_\alpha \nabla_\beta A^\mu - \nabla_\beta \nabla_\alpha A^\mu \quad (5)$$

- f) four-dimensional Ricci-tensor:

$$R_{\nu\beta} = R_{\nu\alpha\beta}^\alpha \quad (6)$$

- g) Einstein's equations:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi T_{\mu\nu} \quad (7)$$

- h) three-dimensional curvature:

$${}^3R_{jkl}^i A^j \equiv D_k D_l A^i - D_l D_k A^i \quad (8)$$

The definition of the three-dimensional Ricci tensor is analogous to the four-dimensional case. Repeated indices are summed.

i) symmetrization:

$$A_{(ij)} \equiv \frac{1}{2}(A_{ij} + A_{ji}) \quad (9)$$

j) anti-symmetrization:

$$A_{[ij]} \equiv \frac{1}{2}(A_{ij} - A_{ji}) \quad (10)$$

2 The incompatibility of quantum cosmology with the Copenhagen interpretation of quantum mechanics

In this section we will set the motivations to study quantum cosmology and show the incompatibility of the Copenhagen interpretation of quantum mechanics with a quantum theory of the universe.

2.1) The motivation for studying quantum cosmology

First, we will present the arguments indicating why a classical model for the universe which maintains at any scale the present laws of physics has probably an initial singularity. For details on this subject, see the book of Ellis and Hawking where it talks about the singularity theorems [1].

Take a timelike four-vector field V^μ with $V^\mu V_\mu = -1$, which may describe the histories of small test particles moving with this velocity or the flow lines of a fluid. We can divide the covariant derivative of this four-velocity field into its irreducible parts:

$$\nabla_\alpha V_\beta = \frac{\theta}{3} h_{\alpha\beta} + \sigma_{\alpha\beta} + w_{\alpha\beta} + V_\alpha A_\beta \quad (11)$$

where

$$\theta = \nabla_\alpha V^\alpha \quad (12)$$

$$\sigma_{\alpha\beta} = h_{(\alpha}^\mu h_{\beta)}^\lambda \nabla_\lambda V_\mu - \frac{1}{3} h_{\alpha\beta} \quad (13)$$

$$w_{\alpha\beta} = h_{[\alpha}^\mu h_{\beta]}^\lambda \nabla_\lambda V_\mu \quad (14)$$

$$A^\alpha = V^\mu \nabla_\mu V^\alpha \equiv \dot{V}^\alpha \quad (15)$$

and $h_{\mu\nu} \equiv g_{\mu\nu} + V_\mu V_\nu$ is the projector onto the surface perpendicular to V^μ at each spacetime point.

These quantities can be interpreted in the following way [5]: the quantity θ is the rate of change of the volume of the fluid represented by the field V^μ , the quantity $\sigma_{\alpha\beta}$ is the

change in form of a constant volume element of the fluid (called the shear tensor), $w_{\alpha\beta}$ is the fluid's rotation, keeping its form and volume constants (called the vorticity tensor), and A^α is its acceleration.

Taking the definition of curvature,

$$R^\mu_{\nu\alpha\beta}V^\nu \equiv \nabla_\alpha \nabla_\beta V^\mu - \nabla_\beta \nabla_\alpha V^\mu \quad (16)$$

summing in μ and α , contracting with V^β , and using equation (11) we obtain, after some manipulation,

$$\dot{\theta} = -\frac{1}{3}\theta^2 - 2\sigma^2 + 2w^2 - R_{\mu\nu}V^\mu V^\nu + \nabla_\alpha A^\alpha \quad (17)$$

where

$$2\sigma^2 \equiv \sigma_{\alpha\beta}\sigma^{\alpha\beta} \quad (18)$$

$$2w^2 \equiv w_{\alpha\beta}w^{\alpha\beta} \quad (19)$$

Equation (17) is called the Raychaudhuri equation. It gives the acceleration of the volume of the fluid. We will assume, for simplicity, that the acceleration of the fluid is zero (a geodesic fluid). For more details on this point we refer again to [1]. The rotation of the fluid gives a positive contribution to the acceleration, in analogy with centrifugal force. However, rotation is linked with closed timelike curves, which violates causality. Therefore, we will set $w^2 = 0$. The first and second term of the right-hand-side (RHS) of equation (17) give a negative contribution to the volume acceleration. The third term is, using Einstein's equations given in the introduction,

$$R_{\mu\nu}V^\mu V^\nu = 8\pi(T_{\mu\nu}V^\mu V^\nu - \frac{T}{2}) \quad (20)$$

which is positive for usual physical fluids. For instance, for a perfect fluid, the RHS of the above equation is $4\pi(\rho + 3p)$ where ρ and p are the energy density and pressure of the fluid, respectively. When the RHS of the above equation is positive or null it is said that the fluid satisfies the *strong energy condition* [1].

Consequently, assuming the hypothesis of the nonexistence of closed timelike curves¹, supposing that Einstein's equations are valid everywhere and that the cosmological fluid satisfies the physically reasonable strong energy condition, equation (17) yields:

$$\dot{\theta} = -\frac{1}{3}\theta^2 - 2\sigma^2 - R_{\mu\nu}V^\mu V^\nu \leq 0 \quad (21)$$

As the volume acceleration is always negative (an expression of the attractive nature of gravity), and knowing that the universe is now expanding, we conclude that the congruence of the timelike curves representing the cosmological fluid has shrunk to zero volume at some finite time in the past. This is the initial cosmological singularity².

¹Recently, a lot of research has been done to investigate if this hypothesis can be proven. The chronology protection conjecture advocates that the laws of physics prevents the existence of closed timelike curves [6, 7].

²A zero volume congruence of curves does not necessarily implies a divergence in the curvature or in the energy density. See reference [8] for a discussion of this point

In order to avoid this pathology we may suppose that, under these extreme situations, general relativity (in a strict sense) is not valid or that the strong energy condition is violated. Some attempts in these directions are theories with non-minimal coupling [9], Weyl geometries [10], change of signature [11], existence of a negative energy scalar field [12] or viscosity effects [13], among others [14].

We will adopt here the position that near strong gravitational fields quantum effects of gravitation become important. We do that for two reasons. First because this is a natural thing to do. Historically, theories that developed singularities were cured by quantum mechanics (like electrodynamics). Also, a world of quantized fields (quantum electro-weak dynamics and quantum chromodynamics) in interaction with an hypothetical fundamental classical gravitational field is inconsistent [8, 15]. Furthermore, the fundamental constants G (Newton's constant), \hbar (Planck's constant) and c (speed of light) yield the fundamental scales where some quantum theory of gravity might be relevant: the Planck scale. They are:

$$\begin{aligned}
 L_{pl} &= \sqrt{\frac{\hbar G}{c^3}} \sim 10^{-33} \text{ cm} \\
 T_{pl} &= \sqrt{\frac{\hbar G}{c^5}} \sim 10^{-45} \text{ s} \\
 M_{pl} &= \sqrt{\frac{\hbar c}{G}} \sim 10^{-5} \text{ g} \\
 \rho_{pl} &= \frac{c^5}{\hbar G^2} \sim 10^{94} \text{ g/cm}^3
 \end{aligned} \tag{22}$$

which are, respectively, the Planck length, the Planck time, the Planck mass and the Planck density.

The second reason for adopting this position is that a theory of quantum gravity when applied to cosmology may also be a theory of initial conditions for the universe. But why is important to have a theory of initial conditions for the universe? This is because the universe we live in is remarkably homogeneous and isotropic, with very small deviations from this highly symmetric state, which are enhanced, in the course of time, by the gravitational interaction (see in this volume the companion article of R. Brandenberger). Clearly, solutions of Einstein's equations with this symmetry are of measure zero; so, why is not the universe inhomogeneous and/or anisotropic? The reader may object by saying that someone who lived in an asymmetric universe could also ask why the universe has this specific kind of inhomogeneity and not another. However, this is a quite different situation. Let us make the following analogy. Suppose there is a couple with 5 children, all born in different years. They can try to calculate the probability of having these specific children out of the totality of genetic possibilities. They will be amazed with the very small result but will correctly reason that this would be the case for every other ensemble of children. Suppose now that these 5 children, without being born in the same gestation (they are not twins), are genetic equal. Five children, born in different gestations, but all genetically identical and, consequently, apart of age, physically identical. The probability for this to happen is as small as any other specific configuration of children. However, the parents will be certainly amazed with such an odd thing and will correctly try to find a doctor

that could explain them why such a bizarre occurrence has happened. In other words, they will naturally suppose, and the doctor too, that there is some deep reason that could explain this phenomenon, some kind of strange behaviour of their reproductive system. The situation is the same in cosmology, and quantum cosmology may be the theory that could explain us the remarkable coincidence that the universe looks like the same in every direction and from every point.

Inflation [16, 17] is an idea that try to explain this coincidence. However, in order for inflation to happen, some special initial conditions are still necessary. As Penrose has pointed out, if we take any inhomogeneous cosmological solution of Einstein's equation and turn it back in time, we will arrive at some initial configuration which is certainly not an initial condition for inflation. Also, the universe may recollapse before inflation takes place. Hence, inflation ameliorate but does not solve the problem. An approach followed by some researchers is to apply quantum cosmology arguments in order to obtain the initial conditions for having inflation.

These are the motivations for studying quantum cosmology, which is the application of quantum gravity to cosmology. However, apart from the problems of quantum gravity itself, which will be discussed in the next section, there is the problem of applying quantum mechanical ideas to a single system as is the universe. In particular, we will now show that the Copenhagen interpretation is not appropriate to quantum cosmology. Let us then make a brief review of this interpretation in the context of non-relativistic quantum mechanics.

2.2) The problem with the Copenhagen interpretation

The postulates of quantum mechanics are³:

- 1) Every state of the system is fixed by a ket $|\Psi(t_0)\rangle$ which belongs to a Hilbert space.
- 2) Every measurable physical quantity is described by an hermitean operator (called an observable) acting in the Hilbert space of the system.
- 3) The only possible results of a measurement of a physical quantity are one of the eigenvalues of the observable associated with it.
- 4) The probability of finding one of these eigenvalues (say, a_n) is given by:

$$p(a_n) = |P_n |\Psi\rangle|^2 = \langle \Psi | P_n | \Psi \rangle \quad (23)$$

where P_n is the projector onto the eigensubspace of the Hilbert space with eigenvalue a_n .

- 5) After a measurement giving the eigenvalue a_n , the state of the system collapses to a new state given by:

$$|\bar{\Psi}\rangle = \frac{P_n |\Psi\rangle}{\sqrt{\langle \Psi | P_n | \Psi \rangle}} \quad (24)$$

³The postulates involving spin and identical particles will not be presented. They are not essential for what follows

6) The evolution of the state of the system is governed by the Schrödinger equation:

$$i\hbar \frac{d|\Psi(t)\rangle}{dt} = \hat{H}(t)|\Psi(t)\rangle \quad (25)$$

where $\hat{H}(t)$ is the hamiltonian operator of the system.

In order to describe statistical mixtures of quantum states, we need another mathematical entity, the density matrix ⁴. A pure state can also be entirely described by the density matrix $\rho \equiv |\Psi\rangle\langle\Psi|$. All equations of the postulates can be written solely in terms of ρ . For instance, the Schrödinger equation can be written as

$$i\hbar \frac{d\rho(t)}{dt} = [\hat{H}(t), \rho(t)] \quad (26)$$

and the probability given in equation (23) is, in terms of ρ :

$$p(a_n) = \text{Tr}(P_n \rho) \quad (27)$$

where Tr means the trace of an operator.

A mixture state can be described by the following density matrix:

$$\rho = \sum_{i=1}^n p_i |\Psi_i\rangle\langle\Psi_i| \quad (28)$$

where p_i is the probability of finding the state $|\Psi_i\rangle$ in the statistical ensemble of states described by the density matrix of equation (27): $\sum_{i=1}^n p_i = 1$. It is a classical statistical distribution of states $|\Psi_i\rangle$ because there is no correlation (represented by off-diagonal terms) among them. In fact, if we want to calculate the probability of finding the eigenvalue a_n of an observable \hat{A} we have, using equations (23) and (27):

$$p(a_n) = \text{Tr}(\rho P_n) = \sum_{i=1}^n p_i \langle\Psi_i|P_n|\Psi_i\rangle \quad (29)$$

which is the sum of the probability of finding the eigenvalue a_n in each state $|\Psi_i\rangle$ multiplied by the probability p_i of finding this state in the statistical ensemble. The reader can easily verify that, if the density matrix ρ has had off-diagonal terms, the probability $p(a_n)$ given above would have been modified with the addition of extra terms representing the quantum interference among the states present in the off-diagonal terms. Hence, a classical statistical mixture of quantum states is necessarily represented by a diagonal density matrix in these states.

Returning to the postulates of quantum mechanics, we would like to make two important remarks. First, we see that time, in spite of being a physical measurable quantity, is not an hermitean operator but a c-number. It is the sole exception of postulate 2. We will discuss this issue in the next section. Second, it seems that there are two laws of evolution for the state vector $|\Psi\rangle$. One is the Schrödinger equation (25) and the other is

⁴A statistical mixture of quantum states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ with weights $|\lambda_1|^2$ and $|\lambda_2|^2$ cannot be described by the state $|\Psi\rangle = \lambda_1 |\Psi_1\rangle + \lambda_2 |\Psi_2\rangle$ due to quantum interference.

the collapse represented by equation (24) which happens when a measurement takes place. This is certainly an odd thing because any measuring apparatus is constituted of atoms and we should expect that its evolution and interaction with the system to be measured should also be described by the Schrödinger equation. Hence, a natural question to ask is: can the Schrödinger equation explain the collapse of the state vector? To answer this question, we will now describe a simple model of what should be a measurement, trying to keep all its essential features. For details, see [18, 19].

Let \hat{S} be the observable that will be measured in a quantum system and $|s\rangle$ its eigenstates: $\hat{S}|s\rangle = s|s\rangle$. Let $|x\rangle$ be the eigenstate of the position operator of some pointer of the apparatus: $\hat{X}|x\rangle = x|x\rangle$. The interaction between the system and apparatus takes place in a finite time interval and during this interval it is much greater than other interactions. Out of this interval, the measured system is isolated from the apparatus, and we can write the state of the two systems as a tensor product of states belonging to their respective Hilbert spaces: $|\Psi\rangle = |\varphi_S\rangle \otimes |\varphi_A\rangle$. The interaction (measurement) will introduce a correlation between these states. If the initial state of the system is an eigenvector of \hat{S} , the measurement will not change it but it must change the state of the apparatus by something proportional to its eigenvalue in order to register this value. Thus, the interaction evolution operator must change the state before the measurement to a state after the measurement in the following way:

$$\hat{U}_I |s\rangle \otimes |x\rangle = |s\rangle \otimes |x + \lambda s\rangle \quad (30)$$

where \hat{U}_I is the interaction evolution operator, and λ is some large coupling constant which realizes the amplification that took place in the apparatus ⁵.

If we now take as initial state the state vector

$$|\Psi_0\rangle = |\varphi_S\rangle \otimes |\varphi_A\rangle \quad (31)$$

with (we are supposing that the eigenvalues s and x are discrete and continuous, respectively):

$$|\varphi_S\rangle = \sum_s c_s |s\rangle \quad (32)$$

and

$$|\varphi_A\rangle = \int f(x) |x\rangle dx \quad (33)$$

we obtain for the final state:

$$|\Psi_F\rangle = \sum_s c_s |s\rangle \otimes |\varphi_A(s)\rangle \quad (34)$$

where $|\varphi_A(s)\rangle = \int f(x) |x + \lambda s\rangle dx = \int f(x - \lambda s) |x\rangle dx$.

Supposing that $f(x)$ is a gaussian centered at $x = 0$ with error Δx (which means that the pointer of the apparatus is at the position $x = 0$ with error Δx) and assuming that

⁵A realization of this interaction can be obtained by taking as interaction hamiltonian the operator $\hat{H}_I = -g(t)\hat{S} \otimes \hat{P}$ where \hat{P} is the observable canonically conjugate to the position \hat{X} of the pointer: $[\hat{X}, \hat{P}] = i\hbar$. If we work in the interaction picture, supposing that $g(t)$ is nonzero only at $-\epsilon < t < \epsilon$ and dominates all other effects at this interval, then we can arrive at equation (30) with $\lambda = \int_{-\epsilon}^{\epsilon} g(t) dt$.

the difference in any pair of eigenvalues s of the operator \hat{S} is much greater than Δx , $\Delta s \gg \Delta x$, (in order for the pointer give a readable result for the measurement) then we can show that:

$$\langle \varphi_A(s) | \varphi_A(s') \rangle = \delta_{s,s'} \quad (35)$$

Therefore, equation (34) means that the final state of the measured system plus apparatus is an orthogonal superposition of states, each containing an eigenvector of \hat{S} with eigenvalue s and a state of the apparatus where the pointer is centered at $x = \lambda s$ (not anymore at $x = 0$ as the initial state).

To be more realistic, let us take into account other degrees of freedom of the apparatus, which is certainly a macroscopic object with many degrees of freedom, and its environment. The transition from the initial to the final state can now be written as:

$$|\Psi_I\rangle = \left(\sum_s c_s |s\rangle \right) \otimes |\varphi_A(r)\rangle \longrightarrow |\Psi_F\rangle = \sum_{s,r'} w_{r,r'}^s c_s |s\rangle \otimes |\varphi_A(s,r')\rangle \quad (36)$$

and $|\varphi_A(r)\rangle = \int f(x) |x\rangle \otimes |r\rangle dx$, $|\varphi_A(s,r)\rangle = \int f(x - \lambda s) |x\rangle \otimes |r\rangle dx$. The variable r represents the extra degrees of freedom and $\langle x', r' | x, r \rangle = \delta(x - x') \delta_{r,r'}$ (for simplicity, we have assumed that the set $\{r\}$ is discrete). The coefficients $w_{r,r'}^s$ are included in order to be as general as possible. As the time evolution coming from the Schrödinger equation preserves the norm than $\sum_{r'} |w_{r,r'}^s|^2 = 1$ which means that the quantum probability to obtain the eigenvalue s continues to be $\sum_{r'} |c_s|^2 |w_{r,r'}^s|^2 = |c_s|^2$.

What should be the density matrix of the observed system plus the measuring apparatus after a real measurement has been performed? We must expect that the final distribution of data should be described by ordinary probability calculus. If it is a real measurement, the different data must be clearly separated events. Therefore, in our example, we must expect that the final density matrix should describe a classical statistical ensemble of states. It should be the tensor product of eigenstates $|s\rangle$ of the measured observable \hat{S} with states of the apparatus describing the position of the pointer dislocated from the initial position by something proportional to the corresponding eigenvalue s , each one of these states appearing in the statistical mixture with probability $|c_s|^2$.

To check if this is true, let us now calculate the density matrix of the final state given in equation (36). As the relevant degree of freedom for the measurement is the position of the pointer, we will calculate the reduced density matrix, which is obtained from the total one by tracing out the irrelevant degrees of freedom r : $\rho_{red} = \sum_r \langle r | \rho | r \rangle$. Everything relative solely to the position of the pointer and to the observed system itself can be calculated with this reduced density matrix [20]. For the final state $|\Psi_F\rangle$ this matrix is:

$$\begin{aligned} \rho_{red} &= \sum_s |c_s|^2 |s\rangle \langle s| \otimes |\varphi_A(s)\rangle \langle \varphi_A(s)| + \\ &+ \sum_{r',s,s':s \neq s'} c_s c_{s'}^* w_{r,r'}^s w_{r,r'}^{*s'} |s\rangle \langle s'| \otimes |\varphi_A(s)\rangle \langle \varphi_A(s')| \end{aligned} \quad (37)$$

The first term in the sum (37) is what we were expecting for a density matrix describing a real measurement. The second term, however, as we have discussed before, describes quantum interference in the data distribution, which is unacceptable for a real measurement. The apparatus, even being a macroscopic system, would be subjected to quantum

interference among its macroscopic states, a situation that has never been observed: no one has ever seen the pointer of an apparatus in a superposition of quantum states like the image in a photograph of superimposed pictures.

The situation does not change if we add a second apparatus that measures the first one and the system [18]. How can we explain this conflict between theoretical description and what is really observed?

The first way out is the Copenhagen interpretation. It says that what is wrong in the theoretical model is the assumption that the macroscopic apparatus can be described by quantum mechanics: it must be described by classical physics. Therefore, it postulates a fundamental division between the quantum world and the classical world, and it is in the last one where observations, experiments, and the consequent knowledge about a quantum system can be acquired. Hence, a quantum system does not have any meaning without a classical world. However, many other questions can be raised: When a system can be considered as macroscopic? What happens in the transition from quantum to classical? What kind of theory is quantum mechanics that has another theory, classical mechanics, as a limit and yet depends on it to have a meaning? Which one is more fundamental? Must the world be described by these two completely different theories? For an excellent discussion about these points, see the book of Omnès [21].

As we can see, this interpretation cannot be applied to quantum cosmology because there is no classical domain in a quantum universe that could give a meaning to the quantum theory. Thus, we must seek other solutions to our problem.

One solution is to think the transition from the initial state before the measurement to the final state after the measurement, described in equation (36), as an splitting of the world into many worlds, each one containing one and only one possible result for the measurement. At each time a measurement takes place, the world is splitted in such a way. No world has knowledge of the other. In each world there is an observer who sees the pointer of the apparatus dislocated by an amount proportional to a particular eigenvalue s . This is the many-worlds interpretation of quantum mechanics [22, 23, 24], developed with the motivation to be applied to the whole universe. This interpretation and some variations of it are frequently used in quantum cosmology. It will be discussed in section 4.

A second solution is to think that the second term in equation (37) which is responsible for the quantum interference, vanishes for almost every macroscopic system. This is not unpalatable because the sum over r involves many degrees of freedom, as it is a macroscopic system, and something like destructive interference may lead this term to be almost zero. In fact, some calculations show that this is indeed the case for the majority of macroscopic systems in nature, with some exceptions that has been subject of intense investigations. This is the decoherence effect [25, 26, 27, 28, 29] and it is so fast that it has never been observed before. It explains how the transition from quantum to classical takes place. The applications of this idea to quantum cosmology will be discussed in sections 5 and 6.

Another possible solution to our problem is to give, like in classical mechanics, an ontological interpretation to quantum mechanics. This means that quantum mechanics should not be interpreted as simply an epistemological theory of given phenomena, but that the processes occurring in the quantum world can be described by deterministic laws, and the physical quantities of a quantum system have a reality independent of any

observation. Therefore, the process of a measurement must be described in a completely different way. In section 6 we will discuss one proposal of this kind of interpretation [30, 31] and its possible application to quantum cosmology.

One common feature of all these alternatives is that, as the classical domain is not put by hand, it must be obtained. In other words, quantum cosmology, when provided with a consistent interpretation of quantum mechanics, must explain why the classical world exists.

After discussing the problems with the application of the Copenhagen interpretation to quantum cosmology, and before applying the possible alternatives to it, let us set up the dynamical equations that a wave function of the universe should satisfy.

3 Canonical quantization of general relativity

As we have explained in the introduction, quantum cosmology is the application of a theory of quantum gravity to the universe as a whole. However, until now, there is no consistent and widely accepted theory of quantum gravity.

There are two basic approaches to quantize gravity. The point of departure is, of course, classical general relativity. One is the covariant approach, which treats the gravitational field as any other field theory. The spacetime metric is splitted in a background or kinematical part, usually taken to be the flat metric $\eta_{\mu\nu}$, and another dynamical part $h_{\mu\nu}$: $g_{\mu\nu} = \eta_{\mu\nu} + Gh_{\mu\nu}$. This splitting is inserted into the general relativity action and the theory is treated perturbatively, making use of the powerful technics of perturbative quantum field theory. The quanta of the field $h_{\mu\nu}$ are viewed as spin-two particles, called gravitons, propagating in the background spacetime, and interacting with itself and with matter. This theory, however, is not renormalizable [32], which means that if we include radiative effects, an infinite number of new parameters must be added and the theory loses its predictive power. Then it was believed that general relativity should be a low energy limit of some more fundamental gravity theory with a better high energy behaviour, exactly like the weak interaction is with respect to the electro-weak interaction. Higher derivative terms were added to the general relativity action. The resulting theory was shown to be renormalizable [33], asymptotically free, but it is not unitary: it does not conserve probability. Other theories with suitable interactions of gravity with matter were developed, like supergravity theories, and they are unitary but still not renormalizable at more than two loops. The last hope in the covariant approach is superstring theory. It is generally believed that the theory is perturbatively finite. However, Gross and Periwal [34] have shown that the whole series of the bosonic string diverges, and they give arguments advocating that this should also happen in superstring theory.

As we have seen, the problems of the covariant approach are very difficult to solve. By solving one problem, one gets another, and it seems that it does not have an end. Perhaps these difficulties are showing us that we should not assume that the spacetime metric can be splitted in a background and a dynamical part. In fact, this is contrary to the spirit of general relativity. One of the motivations of Einstein to construct his theory was to get rid of objects that act on other objects and are not influenced by them. Background metrics are exactly like that. General relativity is a theory with no background metric;

the spacetime metric is known only after the equations of motion are solved. Therefore, we should try to construct a non-perturbative theory of quantum general relativity. Can quantum general relativity make sense if its perturbation expansion is not renormalizable? The answer is affirmative. There are examples of theories that are exactly solvable non-perturbatively but which perturbation expansion is not renormalizable [35]. This leads us to the second approach: quantize gravity by non-perturbative methods. Non-perturbative technics are being developed in superstring theory, but we will stay in the framework of general relativity itself and present the canonical quantization of this theory.

The canonical approach is based on the hamiltonian of general relativity. The idea is to obtain a quantum functional equation for a wave functional, which is analogous to the Schrödinger equation. For historical reasons, this approach is not very popular in other quantum field theories. Some papers have been published with comparisons of this approach with the more usual covariant approach in quantum electrodynamics and other quantum field theories [36, 37, 38]. To construct the hamiltonian of general relativity we must assume that spacetime can be splitted into a family of spacelike hypersurfaces and a timelike direction. It means that we are restricting the topology of the manifold to be of the type: $M^4 = R \otimes M^3$. Hence, we are discarding spacetimes with rotation and with closed timelike curves, in accordance with the assumptions of section 2. Questions about the existence of closed timelike curves cannot be answered within this formalism.

Let us now split the metric into the timelike direction and the spacelike direction.

The spacelike hypersurfaces can be defined by the equations $\phi(x^\mu) = \text{const.}$. Their normals are given by one-forms $\eta = \eta_\mu dx^\mu = \partial_\mu \phi dx^\mu$. As they are spacelike, there is always a timelike coordinate $x^0 = t$ that parametrizes the hypersurfaces yielding $\eta_\mu = -N \delta_\mu^0$. N is a normalization factor, $g^{\mu\nu} \eta_\mu \eta_\nu = -1$, which implies that $g^{00} = -\frac{1}{N^2}$. The projector onto the hypersurfaces is given by $h^{\mu\nu} \equiv g^{\mu\nu} + \eta^\mu \eta^\nu$ whose components are $h^{00} = 0$, $h^{0i} = 0$ and $h^{ij} = g^{ij} + N^2 g^{i0} g^{j0}$. Defining $N^i = g^{i0} N^2$, the components of the contravariant metric are:

$$g^{00} = -\frac{1}{N^2}; \quad g^{0i} = \frac{N^i}{N^2}; \quad g^{ij} = h^{ij} - \frac{N^i N^j}{N^2} \quad (38)$$

We can calculate the inverse covariant metric $g_{\mu\nu}$ yielding the following line element:

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= (N_i N^i - N^2) dt^2 + 2N_i dx^i dt + h_{ij} dx^i dx^j = \\ &= N^2 dt^2 + h_{ij} (N^i dt + dx^i) (N^j dt + dx^j) \end{aligned} \quad (39)$$

where $N_i = h_{ij} N^j$, h_{ij} is the inverse of h^{ij} and it is, by its construction, the intrinsic covariant metric of the spacelike hypersurfaces. Examining equation (39) we can see that $N(t, x^k)$ is the rate of change with respect to the coordinate time t of the proper time of an observer with four-velocity $\eta^\mu(t, x^k)$ at the point (t, x^k) . It is called the lapse function. Also, $N^i(t, x^k)$ is the rate of change with respect to coordinate time t of the shift of the points with the same label x^i when we go from one hypersurface to another. It is called the shift function. It can also be viewed as the projection onto the spacelike hypersurface of the tangent vector $\frac{\partial}{\partial t}$ to the t-time coordinate curves. For more details on this, see reference [39].

Another useful quantity is the extrinsic curvature. It measures how much the 3-dimensional hypersurfaces are curved with respect to the 4-dimensional manifold in which

it is embedded. It does that by comparing the normal vector η_μ at one point with the parallel transported normal vector from a neighbour point to this same point. Precisely, it is defined as follows:

$$K_{\mu\nu} \equiv -h_\mu^\alpha h_\nu^\beta \nabla_{(\alpha} \eta_{\beta)} \quad (40)$$

The relevant components of the extrinsic curvature are:

$$\begin{aligned} K_{ij} &= -N\Gamma_{ij}^0 \\ &= \frac{1}{2N}(2D_{(i}N_{j)} - \partial_t h_{ij}), \end{aligned} \quad (41)$$

Using equations (38), (39) and (41), we obtain for the four-dimensional scalar of curvature:

$$R = R^{(3)} + K^{ki}K_{ki} + K^2 - \frac{2}{N}\partial_t K + \frac{2N^i}{N}\partial_i K - \frac{2}{N}D_k(\partial^k N). \quad (42)$$

The Einstein-Hilbert lagrangian density can be written as:

$$\begin{aligned} \mathcal{L}_E &= \sqrt{-g}R = Nh^{1/2}R \\ &= Nh^{1/2}(R^{(3)} + K_{ij}K^{ij} - K^2) - 2\partial_t(h^{1/2}K) + \\ &\quad + 2\partial_i(h^{1/2}KN^i - h^{1/2}h^{ki}\partial_k N). \end{aligned} \quad (43)$$

There are two total derivatives in this lagrangian density. The total time derivative leads to inconsistencies in the path integral formulation of the quantum theory [40]. Furthermore, with this term, we cannot obtain the gravitational part of Einstein's equations by simply varying the lagrangian density (43) with respect to N , N^i and h_{ij} , and imposing, as usual, that the variations of these quantities on the boundaries are zero. We need to impose as well that the time derivative of the variations of h_{ij} are also zero. For these reasons, we will eliminate this term by taking the modified lagrangian density:

$$\mathcal{L} \equiv \mathcal{L}_E + 2\partial_t(h^{1/2}K) \quad (44)$$

The last term in equation (43) is not important for the lagrangian formalism. In fact it is an arbitrary term because we can add or subtract total spatial derivative terms to the lagrangian without changing the lagrangian equations of motion. However, such terms are crucial for the hamiltonian formalism. For open spaces, like asymptotically flat or anti-de Sitter spacetimes, this term must be chosen judiciously if we want to obtain the correct hamiltonian equations of motion [41]. They also yield the total gravitational energy of such spaces (when it can be defined). They are very relevant in the study of quantum black holes. In the quantum cosmology of a closed universe, however, these terms are zero and can be discarded from the lagrangian. The reader may ask why quantum cosmology does not deal with open universes. First, closed universes are technically simpler and conceptually richer. We hope the reader will be convinced of this assertion by the end of these lectures. Second, there are path integral arguments claiming that open universes are not probable. We will return to this point in section 5. However, there is no convincing argument pointing in this direction. Therefore, we will make a comment every time the presence of such terms can give different results.

Hence, the gravitational lagrangian density will be taken to be:

$$\mathcal{L}[N, N^i, h_{ij}] = N h^{1/2} (R^{(3)} + K^{ij} K_{ij} - K^2). \quad (45)$$

The total lagrangian is evidently given by:

$$L = \int \mathcal{L} d^3x \quad (46)$$

Variation with respect to N , N^i and h_{ij} gives the projections of vacuum Einstein's equations $G_{\mu\nu} \eta^\mu \eta^\nu = 0$, $G_{\mu\nu} \eta^\mu h^\nu_\alpha = 0$ and $G_{\mu\nu} h^\mu_\beta h^\nu_\alpha = 0$, respectively.

Let us now construct the hamiltonian of general relativity. As the lagrangian density (46) does not depend on $\partial_t N$ and on $\partial_t N^i$, their canonical conjugate momenta are zero:

$$\Pi_\mu = \frac{\delta L}{\delta(\partial_0 N^\mu)} \approx 0 \quad (47)$$

where $N^0 \equiv N$.

The symbol \approx means ‘weakly zero’ to remind us that the Poisson brackets of these quantities with other functions of phase space variables may not be equal to zero.

Therefore, general relativity is a theory with constraints and it will be treated with the Dirac formalism [42, 43, 44]. In the language of Dirac, the constraints (47) are called primary constraints.

The canonical momenta conjugate to h^{ij} are given by:

$$\Pi_{ij} = \frac{\delta L}{\delta(\partial_t h^{ij})} = -h^{1/2} (K_{ij} - h_{ij} K) \quad (48)$$

The canonical hamiltonian density \mathcal{H}_c is obtained in the usual way: $\mathcal{H}_c = \Pi_{ij} \partial_t h^{ij} - \mathcal{L}$. After discarding a divergence term (something that is not generally possible for open spacetimes), it yields the following canonical hamiltonian:

$$\begin{aligned} H_c &= \int d^3x \mathcal{H}_c \\ &= \int d^3x (N \mathcal{H} + N_j \mathcal{H}^j), \end{aligned} \quad (49)$$

where

$$\mathcal{H} = G_{ijkl} \Pi^{ij} \Pi^{kl} - h^{1/2} R^{(3)} \quad (50)$$

$$\mathcal{H}^j = -2D_i \Pi^{ij}. \quad (51)$$

and

$$G_{ijkl} = \frac{1}{2} h^{-1/2} (h_{ik} h_{jl} + h_{il} h_{jk} - h_{ij} h_{kl}) \quad (52)$$

which is called the DeWitt metric.

The total hamiltonian density must yield the constraints (47). Thus, it must be given by:

$$\mathcal{H}_T = N \mathcal{H} + N_j \mathcal{H}^j + \lambda^\mu \Pi_\mu \quad (53)$$

where λ^μ are lagrangian multipliers.

For consistency, the primary constraints must be conserved in time: $\dot{\Pi}_\mu = \{\Pi_\mu, \mathcal{H}_T\} = 0$. This implies that the quantities (50) and (51) must be weakly zero:

$$\mathcal{H} = G_{ijkl}\Pi^{ij}\Pi^{kl} - h^{1/2}R^{(3)} \approx 0 \quad (54)$$

$$\mathcal{H}^j = -2D_i\Pi^{ij} \approx 0. \quad (55)$$

They are secondary constraints and are called super-hamiltonian and super-momentum constraints, respectively. Their conservations in time do not lead to any new constraints. As N and N^i have no dynamics and they multiply secondary constraints in the total hamiltonian, they can be viewed as lagrangian multipliers of these constraints, and they can be eliminated from the phase space of the theory [44]. Therefore, the hamiltonian of general relativity is simply given by:

$$H_{GR} = \int d^3x(N\mathcal{H} + N_j\mathcal{H}^j) \quad (56)$$

It can be shown that the secondary constraints have weakly zero Poisson brackets among each other. They are called first class constraints. There is a conjecture of Dirac saying that all first class constraints are generators of gauge transformations. In fact, it can be shown that:

$$\delta h_{ij}(x) = \{h_{ij}(x), \int d^3y \xi^k(y)\mathcal{H}_k(y)\} = D_j\xi_i(x) + D_i\xi_j(x) = \mathcal{L}_\xi h_{ij} \quad (57)$$

$$\delta h_{ij}(x) = \{h_{ij}(x), \int d^3y \zeta(y)\mathcal{H}(y)\} = -2\zeta(x)K_{ij}(x) = \zeta(x)\mathcal{L}_\eta h_{ij} \quad (58)$$

where \mathcal{L}_ξ is the Lie derivative along the infinitesimal spacelike vector ξ and \mathcal{L}_η is the Lie derivative along the direction orthogonal to the spacelike hypersurfaces with metric h_{ij} . The function $\zeta(x)$ is infinitesimal. Analogous results can be obtained for the momenta Π_{ij} . Therefore, the first constraint is the generator of spatial coordinate transformations while the second one is the generator of time reparametrization, which are the gauge transformations of the theory. As can be seen from equation (58), the second constraint is also responsible for the dynamics of the theory.

Variation of H_{GR} with respect to N and N^i yields the constraint equations $\mathcal{H} = 0$ and $\mathcal{H}^i = 0$ which are the vacuum Einstein's equations $G_{\mu\nu}\eta^\mu\eta^\nu = 0$ and $G_{\mu\nu}\eta^\mu h^\nu_\alpha = 0$, respectively. The evolution equation for h^{ij} gives the definition of Π_{ij} shown in equation (48), which, combined with the evolution equation for Π_{ij} , yields the dynamical Einstein's equation $G_{\mu\nu}h^\mu_i h^\nu_j - 2h_{\alpha(j}G_{i)\nu}\eta^\nu\eta^\alpha = 0$.

As we can see, the vacuum Einstein's equations are obtained from a phase space composed of all possible hypersurface metrics $h^{ij}(x)$ and their canonical momentum $\Pi_{ij}(x)$, which means that the configuration space of the theory is composed by all possible $h^{ij}(x)$. A particular spacetime solution of Einstein's equations can be viewed as a trajectory in the space of all $h^{ij}(x)$. By making an analogy with the dynamics of particles, $h^{ij}(x) \longrightarrow x^i$ and $\Pi^{ij}(x) \longrightarrow p^i$, we can interpret the first term in the constraint (50) as a kinetic term, $G_{ijkl}(h_{ij})$ given in equation (52) playing the role of a metric in the space of metrics (like a $g_{ij}(x)$ in the case of particles), and the second term as a potential energy.

As can be seen from equation (56), the hamiltonian of general relativity is numerically zero because it is a combination of constraints. However, if we were considering open spaces, the spatial surface terms I have discarded may appear and the total hamiltonian is not anymore numerically zero. In fact, for asymptotically flat spacetimes for instance, these surface terms yield their total energy.

Zero hamiltonians are characteristic of time-reparametrization invariant theories, as is general relativity. Take as an example an action describing the classical dynamics of a system of particles:

$$S = \int L(x^i, \frac{dx^i}{dt}, t) dt \quad (59)$$

Let us define a new parameter τ and treat time as a new coordinate depending on τ . The new action is:

$$S = \int L(x^i, \frac{\dot{x}^i}{\dot{t}}, t) \dot{t} d\tau = \int \bar{L}(t, x^i, \dot{t}, \dot{x}^i) \quad (60)$$

where the dot means derivative on τ .

This action is now invariant by reparametrizations $\tau' = \tau'(\tau)$. Let us calculate its hamiltonian. The canonical momenta are:

$$\pi_i = \frac{\partial \bar{L}}{\partial \dot{x}^i} = \frac{\partial L}{\partial (\frac{dx^i}{dt})} \equiv p_i \quad (61)$$

$$\pi_0 = \frac{\partial \bar{L}}{\partial \dot{t}} = -H(x^i, p_j, t) = -H(x^i, \pi_j, t) \quad (62)$$

where $H(x^i, p_j, t)$ denotes the original hamiltonian of the system. Equation (62) is a constraint equation; no time derivative of phase space variables appear in it:

$$\pi_0 + H(x^i, \pi_j, t) \approx 0 \quad (63)$$

The canonical hamiltonian is easily calculated and it is:

$$\bar{H}(t, x^i, \pi_0, \pi_i) = \dot{t}[\pi_0 + H(x^i, \pi_j, t)] \quad (64)$$

which is zero due to the constraint (63).

The total hamiltonian will be given by:

$$H_T = N[\pi_0 + H(x^i, \pi_j, t)] \quad (65)$$

where N is a lagrangian multiplier. This hamiltonian is also zero.

The reader can verify that this hamiltonian gives the correct equations of motion, and that the constraint (63) generates the gauge transformation linked with the reparametrization invariance of the theory. It is the analog of the constraint (54).

Let us return to general relativity and try to quantize the theory. We will work in the h_{ij} representation. The rules for quantization are:

i) Transform phase space variables into operators acting on functionals of h_{ij} and t , $\Psi[h_{ij}, t]$.

ii) Poisson brackets turn into comutators. In particular:

$$\{h_{ij}(x), \Pi^{kl}(x')\} \longrightarrow \frac{1}{i\hbar}[h_{ij}(x), \Pi^{kl}(x')] \quad (66)$$

This means that we can write:

$$\hat{\Pi}^{ij} = -i\hbar \frac{\delta}{\delta h_{ij}}, \quad (67)$$

iii) The wave function $\Psi[h_{ij}, t]$ must satisfy the Schrödinger-like functional equation:

$$i\hbar \frac{\partial \Psi(h_{ij}, t)}{\partial t} = \hat{H}_{GR} \Psi(h_{ij}, t) \quad (68)$$

where \hat{H}_{GR} is the operator coming from the classical hamiltonian (56).

What will be the quantum versions of the constraint equations (54) and (55)? They cannot turn into operator equations due to rule (ii). In fact, if we demand that equations (54) and (55) are operators identities then all comutators with them would be zero. But not all Poisson brackets involving the constraints (54) and (55) are zero and this would be a contradiction with rule (ii). But constraints (54) and (55) are first class constraints and there will be no contradiction if we impose them, as Dirac suggests, as conditions on the wave function $\Psi[h_{ij}, t]$:

$$\hat{H} \Psi(h_{ij}, t) = 0 \quad (69)$$

$$\hat{H}^k \Psi(h_{ij}, t) = 0. \quad (70)$$

If equations (69) and (70) are correct, then the right-hand-side of equation (68) is zero and it implies that Ψ does not depend on t .

Let us go back to our previous example of a time reparametrization invariant action of a system of classical particles to understand what is going on.

If we follow the Dirac rules in this example, the Schrödinger equation will simply imply that the wave function does not depend on τ . Imposing the constraint equation (63) as a condition on the wave function,

$$[\hat{\pi}_0 + \hat{H}(x^i, \pi_j, t)] \Psi(t, x^i) = 0 \quad (71)$$

gives the original Schrödinger equation of the system.

As the constraint (54) is the analog of the constraint (63), we expect that equation (69) contains the dynamics of the wave function. Let us write explicitly equations (69) and (70):

$$2iD_j \frac{\delta \Psi(h^{ij})}{\delta h^{ij}} = 0 \quad (72)$$

$$(\hbar^2 G^{ijkl} \frac{\delta}{\delta h^{ij}} \frac{\delta}{\delta h^{kl}} + h^{1/2} R^{(3)}) \Psi(h^{ij}) = 0. \quad (73)$$

(we have set $\hbar = 1$).

The first equation has a simple interpretation. Make an infinitesimal spatial coordinate transformation

$$x^i \rightarrow x^i - \xi^i \quad (74)$$

The spacelike metric changes in the following way:

$$h_{ij} \rightarrow h_{ij} + 2D_{(i}\xi_{j)}. \quad (75)$$

The new wave function may be expanded yielding:

$$\psi[h_{ij} + 2D_{(i}\xi_{j)}] = \Psi[h_{ij}] + \int d^3x 2D_{(i}\xi_{j)} \frac{\delta\Psi}{\delta h_{ij}} \quad (76)$$

Integrating by parts the second term of the right-hand side of the preceding equation, and as we are supposing that the spacelike hypersurfaces are closed, we obtain the following expression for the change in Ψ :

$$\delta\Psi = - \int d^3x \xi_j D_i \left(\frac{\delta\Psi}{\delta h_{ij}} \right) = 0, \quad (77)$$

where we have used equation (72). It means that the value of the wave function does not change if the spacelike metric changes by a coordinate transformation. Therefore, equation (72) implies that the wave function is a functional of the equivalence class of metrics which describe the same geometry, not of one particular metric. It is a functional defined on the space of all spacelike geometries, not on the space of all spacelike metrics. The space of all three-dimensional spacelike geometries is called superspace. This is the quantum version of the meaning of the constraint (55), which classically was interpreted as the generator of spacelike coordinate transformations.

Let us turn to the equation (73), which is called the Wheeler-DeWitt equation [4]. This equation is the analog of equation (71) for particle dynamics. We should expect that the dynamics of the wave function be contained in it. Like in equation (71), there should exist one momentum which is canonically conjugate to the time in which the quantum dynamics takes place. In equation (71) this particular momentum is easily distinguishable from the others because it appears linearly in this equation, while the others appear quadratically. However, in equation (73), there is no momentum which appears linearly; all of them appear quadratically. Hence, where is time? (Once again it should be reminded that if we were working with open spaces, the total hamiltonian would have had extra surface terms and the Schrödinger equation of the problem would no longer be trivial: *the wave function would depend on time*).

There are some proposals of solution to this problem, which is called the issue of time. We will now expose some of them:

i) The DeWitt metric (52) is a 6×6 matrix per space point and it can be shown that it has signature $(-, +, +, +, +, +)$ [4]. The minus sign is related to the square root of the determinant of the spacelike metric [4, 45, 46], \sqrt{h} . Thus, it seems that we should identify this quantity with time. However, \sqrt{h} is the volume of the spacelike hypersurfaces. Does it mean that if the universe recolapses time will go backwards? Quite unpalusible. Furthermore, as the DeWitt metric has a Lorentzian signature, the Wheeler-DeWitt equation (73) is like a Klein-Gordon equation with a variable ‘mass’ term, $R^{(3)}(h^{ij})$, which depends on the ‘time’ \sqrt{h} . Consequently, if we want to give some kind of probabilistic interpretation to Ψ , we will have to face all the problems with negative probabilities which are

characteristic of this type of equation. The presence of the variable ‘mass’ term turns this problem difficult to solve [47].

In quantum field theory, this problem is solved by second quantizing the Klein-Gordon field. This field operator is expanded in creation and annihilation operators of spin zero particles. The vacuum state is the state with no particles. If this quantum field is submitted to a time variable potential energy or if it is embedded in a time variable curved background, then spin zero particles are created out of the vacuum.

For the Wheeler-DeWitt equation, this procedure would lead us to a third quantization of gravity by quantizing the wave function itself [48, 49]. The particles are now universes that can be created by the action of creation operators which are obtained by an expansion of the wave function, which is now an operator. The vacuum state is the real nothing, the absence of matter *and* spacetime. As the DeWitt metric (52) as well as $R^{(3)}(h^{ij})$ depends on $\sqrt{\hbar}$, which is considered here as ‘time’, then this quantum wave function is like a quantum scalar field propagating in a time variable curved background and submitted to a time variable potential energy. Thus, universes can be spontaneously created from nothing! This is a very exotic and attractive picture. Note that within this picture, it may be possible to explain why the constants of nature have the values we measure of them [50, 51].

ii) We could try to find some variables where the Wheeler-DeWitt equation (73) has the form of equation (71). In fact, this is possible but only implicitly [47, 46]. The variable that plays the role of time is the trace of Π_{ij} which is proportional to $-\theta$ defined in equation (12). It is a good choice because it is a monotonically increasing function of time whenever the dominant energy condition is satisfied (see equation (21)).

iii) The fact that it is not easy to find what should play the role of time in the Wheeler-DeWitt equation simply means that there is no time in quantum gravity [52, 53]. In fact, the analogy with the quantum mechanics of particles via the time reparametrization invariant action (62) is not appropriate. One should take the Jacobi action

$$S = \int d\tau \sqrt{F_E T} \tag{78}$$

where $F_E = E - V$ and $T = \frac{1}{2} \sum_{i=1}^n m_i \frac{dx^i}{d\tau} \frac{dx^i}{d\tau}$. This is the appropriate action when a closed conservative system is studied. The conserved energy is E , and V and T are the potential and kinetic energy of the system. This action yields the Newton equations of motion if a suitable choice of the parameter τ is made such that $T = F_E$.

The hamiltonian can be calculated in the same way as before and it turns out to be proportional to the following constraint:

$$\frac{1}{2} \sum_{i=1}^n \frac{p^i p^i}{m_i} - F_E \approx 0 \tag{79}$$

Following the Dirac quantization scheme, this constraint yields the following quantum

equation:

$$\frac{1}{2} \left(\sum_{i=1}^n \frac{\hat{p}^i \hat{p}^i}{m_i} + V \right) \Psi(x^i) = E \Psi(x^i) \quad (80)$$

which is the time independent Schrödinger equation.

This is the correct analogous equation to the Wheeler-DeWitt equation (73) because it is also quadratic in all momenta. Consequently, we should consider the Wheeler-DeWitt equation as a time-independent Schrödinger equation.

How can we physically justify that? First note that time appears in quantum mechanics as an external parameter. If we want to describe an open system, like an ensemble of black holes we will need open spaces, like asymptotically flat spaces. Consequently, the hamiltonian of general relativity will no longer be zero and the wave function will depend on time, as reminded before. This time comes from the asymptotic structure of such spaces. Thus everything is coherent; time appears because there is an external place where it can come from: the asymptotic structure. However, for closed spaces, there is no place where it can come from. We are quantizing everything, nothing is left. Furthermore, as we have commented before, space geometry is like position in ordinary particle mechanics while spacetime geometry is like a trajectory. As trajectories have no sense in the quantum mechanics of particles, only instantaneous positions have, we can conclude that spacetime has no meaning in quantum gravity, only space geometries have. Hence, time has no sense at the Planck scale. Therefore, it is quite natural that the Wheeler-DeWitt equation of closed spaces be time independent. It is a time independent Schrödinger equation for zero energy, as it should be!

The reader may object that we should expect that the gravitational field should behave like a massless spin-two field and have two degrees of freedom, like in the weak field limit. However, nothing can assure us that this conclusion can be extended to full quantum gravity.

How time can be recovered? It can be recovered only at the semi-classical limit, where geometry becomes classical and spacetime has a sense. To show this, let us take the semi-classical limit of the Wheeler-DeWitt equation, now taking matter into account [54, 55, 56]:

$$\left[\frac{\hbar^2}{2M} G_{ijkl} \frac{\delta}{\delta h_{ij}} \frac{\delta}{\delta h_{kl}} + M h^{1/2} R^{(3)} + \hat{H}_m(h_{ij}, \phi, \pi_\phi) \right] \Psi(h_{ij}, \phi) = 0 \quad (81)$$

where $M \equiv \frac{c^2}{32\pi G}$.

The reader may convince himself of the last term of equation (81) by adding to the lagrangian density of general relativity (46), for example, the lagrangian density of a minimally coupled scalar field and follow the same steps we have followed to arrive at equation (73).

As we are interested in the limit of classical gravity and quantum matter, we will consider the limit where G is small (M is large) as compared with some combination of Planck's constant and the coupling constants of matter.

We write the wave function as:

$$\Psi = e^{iS/\hbar} \quad (82)$$

and expand S in the form $S = MS_0 + S_1 + M^{-1}S_2 \dots$

In the highest order M^2 we obtain:

$$\left(\frac{\delta S_0}{\delta \phi}\right)^2 = 0 \quad (83)$$

This means that S_0 depends only on the metric.

In the next order M^1 we have:

$$\frac{1}{2}(G_{ijkl} \frac{\delta S_0}{\delta h_{ij}} \frac{\delta S_0}{\delta h_{kl}} + h^{1/2} R^{(3)}) = 0 \quad (84)$$

This is the Hamilton-Jacobi equation for vacuum general relativity and they are equivalent to the vacuum Einstein's equations in the sense that once a solution S_0 of this equation is found, the spatial geometry can be integrated to yield the spacetime geometry.

In order M^0 we have:

$$G_{ijkl} \frac{\delta S_0}{\delta h_{ij}} \frac{\delta S_1}{\delta h_{kl}} - \frac{i\hbar}{2} G_{ijkl} \frac{\delta^2 S_0}{\delta h_{ij} \delta h_{kl}} + \frac{1}{2\sqrt{h}} \left(\frac{\delta S_1}{\delta \phi}\right)^2 - \frac{i\hbar}{2\sqrt{h}} \frac{\delta^2 S_1}{\delta \phi^2} = 0 \quad (85)$$

To understand this equation, we define a new functional

$$f \equiv D(h^{ij}) e^{iS_1/\hbar} \quad (86)$$

where the functional $D(h^{ij})$ satisfies the equation:

$$G_{ijkl} \frac{\delta S_0}{\delta h_{ij}} \frac{\delta D}{\delta h_{kl}} - \frac{1}{2} G_{ijkl} \frac{\delta S_0}{\delta h_{ij}} \frac{\delta S_0}{\delta h_{kl}} D = 0 \quad (87)$$

In terms of f , equation (85) can be written as:

$$i\hbar G_{ijkl} \frac{\delta S_0}{\delta h_{ij}} \frac{\delta f}{\delta h_{kl}} \equiv i\hbar \frac{\delta f}{\delta \tau(x)} = \hat{H}_m f \quad (88)$$

This is a Tomonaga-Schwinger equation [57] describing a quantum scalar field propagating in a curved *classical* background which is the solution of equation (84). The many fingered time $\tau(x)$ is the time associated with observers with four velocity η^μ which are orthogonal to the surfaces $S_0 = \text{const.}$ at each spacetime point. It is the natural definition of time because $\frac{\delta S_0}{\delta h_{ij}}$ is just the momentum conjugate to the classical solution h^{ij} , which is linked to the Lie derivative of h^{ij} with respect to the normal vector to the hypersurfaces where h^{ij} is defined. Therefore, time appears when spacetime becomes classical, as we have seen⁶.

One could also try to compute the back reaction of the quantum field ϕ into equation (84) in order to obtain the semiclassical Einstein's equations:

$$G_{\mu\nu} = - \langle T_{\mu\nu} \rangle \quad (89)$$

This problem is treated in references [54, 55].

In this section we have obtained the Schrödinger-like equation for the quantum wave function which is, in the presence of matter, the Wheeler-DeWitt equation (81). We have seen that there is no time in it. In the next section, we will try to interpret this wave function and try to extract physical information from it.

⁶See, however, Ref. [56] for a discussion on conditions for this time be well defined.

4 Predictions from the wave function of the universe

As we have emphasized in section 2, we need a new interpretation of quantum mechanics that can be applied consistently to the wave function of the universe, which is a solution of the Wheeler-DeWitt equation (81).

In quantum mechanics, the sole sentence we can affirm with certainty, which is independent of probabilities, is that the measurement of some observable \hat{B} of a quantum system in a state which is an eigenstate of this observable corresponding to the eigenvalue b , yields the value b . This state, when expressed on the basis of eigenvectors of \hat{B} , will have only one component different from zero: the one associated with the eigenvalue b ⁷. Therefore, we could try to find the operators from which a solution of the Wheeler-DeWitt equation is an eigenfunction. It is evident that it is a very difficult task, perhaps with no solution.

We could try to relax this eigenstate restriction, and create the notion of an approximate eigenstate, which is defined as a state whose wave function has a sharp finite peak at one of the eigenvalues of \hat{B} and smoothly goes to zero outside it. This brings us to a version of the many-worlds interpretation of quantum mechanics cited in section 2. In this version [58, 59], the predictions of quantum mechanics are described in terms of precluded regions. The values of some observable for which the wave function is small, *not necessarily zero*, are *impossible* to be obtained. Note that the usual interpretation of quantum mechanics says that small regions of the wave function are not precluded or impossible; they only have a small probability to occur. However, we cannot talk about probabilities in the quantum mechanics of an individual system. Hence, within this alternative interpretation which attempts to be applied to individual systems, small regions will be treated as precluded regions. Geroch shows in his paper [58] how some known predictions of quantum mechanics can be obtained with the use of this notion of precluded regions. He also emphasizes that if there is some prediction of quantum mechanics that cannot be said in terms of precluded regions, this interpretation must be discarded.

If the individual quantum system is divided into many identical quantum systems, we should expect that the old interpretation in terms of probabilities could be obtained. This can be shown in the following way (for details, see Ref. [59]): suppose we have an observable \hat{S} which, for simplicity, has a discrete spectrum. Its eigenvalues are s_i with respective eigenstates $|i\rangle$. Suppose also that the state of the individual system $|\Psi\rangle$ can be written as a tensor product $|\chi\rangle \otimes |\varphi\rangle$ where the state $|\varphi\rangle$ is also a tensor product of N identical subsystems:

$$|\varphi\rangle = |\phi_1\rangle \otimes \dots \otimes |\phi_j\rangle \otimes \dots \otimes |\phi_N\rangle \quad (90)$$

The identical subsystems are in identical states which can be expanded in the basis of eigenvectors of \hat{S} as:

$$|\phi_j\rangle = |\phi\rangle = \sum_i c_i |i\rangle \quad (91)$$

⁷In the case of a continuous spectrum, the associated wave function will be a Dirac delta. For instance, let \hat{B} be the position operator. Its ‘eigenstate’ with eigenvalue x' is $|x'\rangle$. The associated ‘eigenfunction’ (in quotes because it is not square integrable) is $\langle x|x'\rangle = \delta(x - x')$.

The first equality expresses that the $|\phi_j\rangle$ states are identical. The index j is written just to remind us that the states $|\phi_j\rangle$ belong to different Hilbert subspaces.

The states $|\phi_j\rangle$ are normalized which means that:

$$\sum_i |c_i|^2 = \sum_i |\langle \phi | i \rangle|^2 = 1 \quad (92)$$

We will be interested in the part of the total quantum individual system $|\varphi\rangle$ given in equation (90). Let an observer measure the observable \hat{S} on each of the identical states $|\phi_j\rangle$. He will certainly obtain one of the eigenvalues s_i . Let us define the relative frequency operator as:

$$\hat{f}(s_l) \equiv \sum_{i_1 \dots i_j \dots i_N} |i_1\rangle \otimes \dots \otimes |i_j\rangle \otimes \dots \otimes |i_N\rangle \frac{\sum_{j=1}^N \delta_{l i_j}}{N} \langle i_N | \otimes \dots \otimes \langle i_j | \otimes \dots \otimes \langle i_1 | \quad (93)$$

where the the sum over each i_j is performed in order to cover all possible eigenvectors of \hat{S} at each subspace labelled by the index j .

The eigenvectors of this operator are $|i_1\rangle \otimes \dots \otimes |i_N\rangle$ (one for each sequence of possible measurement results $\{s_{i_1} \dots s_{i_N}\}$):

$$\hat{f}(s_l) |i_1\rangle \otimes \dots \otimes |i_N\rangle = \frac{\sum_{j=1}^N \delta_{l i_j}}{N} |i_1\rangle \otimes \dots \otimes |i_N\rangle \quad (94)$$

As we can see from equation (94), the eigenvalues of the relative frequency operator $\hat{f}(s_l)$ is, as the name indicates, the relative frequency in which the particular eigenvalue s_l appears in the sequence $\{s_{i_1} \dots s_{i_N}\}$ corresponding to the eigenvector $|i_1\rangle \otimes \dots \otimes |i_N\rangle$.

Let us now calculate the norm of the ket $\hat{f}(s_l)|\varphi\rangle - |c_l|^2|\varphi\rangle$. This will be given by:

$$|\hat{f}|\varphi\rangle - |c_l|^2|\varphi\rangle|^2 = \langle \varphi | \hat{f}^2 | \varphi \rangle - 2|c_l|^2 \langle \varphi | \hat{f} | \varphi \rangle + |c_l|^4 \quad (95)$$

First note that:

$$\hat{f}^2 = \sum_{i_1 \dots i_j \dots i_N} |i_1\rangle \otimes \dots \otimes |i_N\rangle \frac{\sum_{j=1}^N \delta_{l i_j}}{N} \frac{\sum_{k=1}^N \delta_{l i_k}}{N} \langle i_N | \otimes \dots \otimes \langle i_1 | \quad (96)$$

The second term in the right-hand-side of equation (95) is proportional to:

$$\begin{aligned} \langle \varphi | \hat{f} | \varphi \rangle &= \sum_{i_1} |\langle \phi_1 | i_1 \rangle|^2 \dots \sum_{i_N} |\langle \phi_N | i_N \rangle|^2 (\delta_{l i_1} + \dots + \delta_{l i_N}) \frac{1}{N} \\ &= |\langle \phi | l \rangle|^2 = |c_l|^2 \end{aligned} \quad (97)$$

where we have used equations (90), (91), (92) and (93).

The first term is proportional to:

$$\begin{aligned} \langle \varphi | \hat{f}^2 | \varphi \rangle &= \sum_{i_1} |\langle \phi_1 | i_1 \rangle|^2 \dots \sum_{i_N} |\langle \phi_N | i_N \rangle|^2 \cdot \\ &\quad \cdot (\delta_{l i_1} \delta_{l i_1} + \dots + \delta_{l i_N} \delta_{l i_N} + 2\delta_{l i_1} \delta_{l i_2} + 2\delta_{l i_1} \delta_{l i_3} \dots) \frac{1}{N^2} = \\ &= \frac{1}{N^2} (N|c_l|^2 + 2 \frac{N(N-1)}{2} |c_l|^4) \end{aligned} \quad (98)$$

where we have used equation (96).

Substituting equations (97) and (98) into equation (95), we obtain:

$$|\hat{f}|\varphi\rangle - |c_l|^2|\varphi\rangle|^2 = \frac{1}{N}(|c_l|^2 - |c_l|^4) \quad (99)$$

In the limit where N goes to infinity, the above norm is zero which means that $|\varphi\rangle$ is an eigenstate of the operator $\hat{f}(s_l)$ with eigenvalue $|c_l|^2$. If N is very large but not infinity, the total wave function $|\varphi\rangle$ in the representation of eigenstates of $\hat{f}(s_l)$ will be sharply peaked around the value $|c_l|^2$. If we use the interpretation that a peak in the wave function is a prediction, we can say that the relative frequency in which the particular eigenvalue s_l is found in a very large sequence of outcomes of measurements of \hat{S} on each of the identical states $|\phi\rangle$, which is nothing but the probability of finding the eigenvalue s_l in a measurement of \hat{S} in the state $|\phi\rangle$, is exactly equal to $|c_l|^2$. Therefore, we recover the usual probabilistic interpretation of quantum mechanics. Note that, in practice, we never make an infinite number of measurements in order to test this probabilistic interpretation.

This result is very attractive and some people clame that this kind of interpretation is more fundamental then the usual one because probabilities are obtained, not postulated⁸.

Adopting this new interpretation, we need to find peaks in the solutions of the Wheeler-DeWitt equation in order to make predictions. However, such solutions are very difficult to find. Furthermore, it is hard to extract physical information from them, due to the absence of the notion of time at the Planck scale, as pointed out in the last section. Can we find other quantum functionals which are more likely to have peaks and which are easier to extract physical information? To answer this question, let us return to ordinary quantum mechanics. There we can construct phase space functions from operators which are functions of \hat{X} and \hat{P} by the so called Weyl-Wigner formalism. In this formalism, a correspondence between an operator \hat{A} which is expressible in terms of coordinate and momentum operators and a function on the phase space of the theory $A(x, p)$ is proposed as follows:

$$A(x, p) \equiv \frac{1}{\hbar^f} \text{Tr}\{\hat{A} \int dy \int dv \exp \frac{i}{\hbar} [(\hat{P} - p) \cdot y + (\hat{X} - x) \cdot v]\} \quad (100)$$

where f is the number of degrees of freedom of the system.

The inverse relation is given by:

$$\hat{A} = \frac{1}{\hbar^{2f}} \int dx \int dp A(x, p) \int dy \int dv \exp \frac{i}{\hbar} [(\hat{P} - p) \cdot y + (\hat{X} - x) \cdot v] \quad (101)$$

Note that $A(x, p)$ depends on \hbar and it is a quantum function.

Other correspondences could be defined. However, these alternatives are unsatisfactory for studying the classical limit. See Ref. [61] for more details on this.

One interesting phase space function is the Wigner function, which is the Weyl-Wigner transform of the density matrix $|\Psi\rangle\langle\Psi|$. It is given by:

$$F(x, p) = \int du \Psi^*(x - \frac{\hbar}{2}u) \Psi(x + \frac{\hbar}{2}u) \exp(-ipu) \quad (102)$$

⁸However, objections to this proof have already been made [60].

It satisfies the following properties:

$$\begin{aligned}\int dp F(x, p) &= |\Psi(x)|^2 \\ \int dx F(x, p) &= |\Psi(p)|^2\end{aligned}\tag{103}$$

The idea proposed in Ref. [62] is to find peaks of the wave function of the universe in the phase space of general relativity by looking for peaks of its corresponding Wigner function of the theory. This idea is motivated by the following reasoning: a semi-classical wave function, in ordinary quantum mechanics, is written in WKB form as:

$$\psi_W(x, t) = A(x, t) \exp\left[\frac{i}{\hbar} S(x, t)\right]\tag{104}$$

where A is a slowly varying function of x and S satisfies the classical Hamilton-Jacobi equation

$$-\frac{\partial S(x, t)}{\partial t} = H(x, p = \frac{\partial S(x, t)}{\partial x})\tag{105}$$

This equation is obtained by inserting the wave function (104) into the Schrödinger equation of the system and keeping only the order- \hbar^0 term⁹.

If we insert this wave function in the Wigner function (102), in order- \hbar^0 , the result obtained in Ref. [62] (which is *not* true, as will see later on) is:

$$F(x, p, t) = |C(x)|^2 \delta\left(p - \frac{\partial S(x, t)}{\partial x}\right)\tag{106}$$

Therefore, WKB wave functions have a peak on the first integral of the equations of motion $p = \frac{\partial S(x, t)}{\partial x}$. If we accept this as a prediction, then f classical solutions are selected out of the $2f$ possible classical solutions of the theory. Furthermore, the prefactor $|C(x)|^2$ can be used as a probability measure on this set of trajectories, as we will see later on. The proposal of Ref. [62] is to follow an analogous procedure: one takes the WKB wave function of the universe and calculates its Wigner functional defined on the phase space of general relativity in order to find the most probable classical cosmological solution of the Einstein's equations.

Let us see an application of this program to a simple example, which is extracted from Ref. [63]. Here we will introduce the notion of a minisuperspace model. The Wheeler-DeWitt equation (81) is a very complicate functional differential equation, which is equivalent to an intricate system of partial differential equations, one for each space point x^i . To solve this equation is evidently very difficult. One usually simplifies it by freezing the degrees of freedom of gravity by reducing the superspace to a minisuperspace where only a finite amount of degrees of freedom are still available.

More precisely, expand the spacelike metric and its conjugate momentum in some complete set f_n :

$$h^{ij}(x, t) = h_{(0)}^{ij}(t) + \sum_{n=1}^{\infty} h_{(n)}^{ij}(t) f_n(x)\tag{107}$$

$$\Pi_{ij}(x, t) = \Pi_{ij}^{(0)}(t) + \sum_{n=1}^{\infty} \Pi_{ij}^{(n)}(t) f_n(x)\tag{108}$$

⁹To show this, is necessary that the classical hamiltonian be a quadratic function of p .

A minisuperspace is the set of all spacelike geometries where all but a set of the $h_{(n)}^{ij}(t)$ and the corresponding $\Pi_{ij}^{(n)}(t)$ are put identically to zero.

Evidently, this procedure violate the uncertainty principle. However, we expect that the quantization of these minisuperspace models retains many of the qualitative features of the full quantum theory, which are easier to study in this simplified model. For more details on minisuperspace models, see Refs. [64, 65, 82].

The minisuperspace model we will discuss was developed in the context of a theory in which gravity is non-minimally coupled to electromagnetism [9], the Lagrangian being given by a combination of Einstein's and Maxwell's theory plus an interacting (non-minimal) term:

$$\mathcal{L} = \sqrt{-g} \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{k} R + \sigma R W_\mu W_\nu g^{\mu\nu} \right] + \partial_t \left[2h^{1/2} K \left(\frac{1}{k} + \sigma W_\mu W_\nu g^{\mu\nu} \right) \right] \quad (109)$$

where σ is a dimensionless positive coupling constant, W_μ is the vector potential and $F_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu$. The surface term appearing in the lagrangian (109) is a generalization, due to the non-minimal coupling, of the one added in general relativity in equation (44).

The field equations of this theory are:

$$(1 + \sigma W^2) G_{\mu\nu} = -\frac{1}{2} E_{\mu\nu} + \sigma \square(W^2) g_{\mu\nu} - \sigma R W_\mu W_\nu - \sigma \nabla_\mu \nabla_\nu (W^2) \quad (110)$$

$$\nabla_\nu F^{\mu\nu} = 2\sigma R W^\mu \quad (111)$$

where $W^2 = g^{\mu\nu} W_\mu W_\nu$, \square is the covariant Laplacian operator, and $E_{\mu\nu} := F_{\mu\alpha} F^\alpha_\nu + \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}$. (We have put $k = 1$).

Our minisuperspace model is characterized by the following ansatz:

$$\begin{cases} ds^2 &= -N^2(t) dt^2 + a^2(t) d\Omega_3^2 \\ W_\mu &= (\psi(t), 0, 0, 0,) \end{cases} \quad (112)$$

The four-metric is of Robertson-Walker form, where $d\Omega_3^2$ is the metric on the spatial sections with constant positive or negative curvature ε ($\varepsilon = +1$ or $\varepsilon = -1$ respectively). The topology of these sections is considered to be closed.

With these assumptions it follows that $F_{\mu\nu} = 0 = E_{\mu\nu}$ and

$$W^2 = g^{\mu\nu} W_\mu W_\nu = -\psi^2/N^2 =: -\phi^2 \quad (113)$$

Defining $\beta := (1 - \sigma\phi^2)$, and substituting σ , ϕ , a and N into the field equations (110) we obtain, after some manipulation:

$$\begin{cases} \frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \varepsilon \frac{N^2}{a^2} - \frac{\dot{a}}{a} \frac{\dot{N}}{N} &= 0 \\ \frac{\ddot{\beta}}{\beta} + 3 \frac{\dot{\beta}}{a} - \frac{3\dot{a}}{a} \frac{\dot{N}}{N} - \frac{\dot{\beta}}{\beta} \frac{\dot{N}}{N} &= 0 \\ \frac{\ddot{\beta}}{\beta} + 3 \frac{\dot{\beta}}{\beta} \frac{\dot{a}}{a} - \frac{\dot{\beta}}{\beta} \frac{\dot{N}}{N} &= 0 \end{cases} \quad (114)$$

From the above equations we obtain the following constraint (no second-order time derivative appears):

$$\dot{a}^2 + a\dot{a} \frac{\dot{\beta}}{\beta} + \varepsilon N^2 = 0 \quad (115)$$

We define the associated minisuperspace action substituting the restriction (112) directly into the lagrangian (109), yielding the action:

$$S := \int dt \mathcal{L}(a, \beta, N) \quad (116)$$

in which

$$\mathcal{L} = \frac{1}{2} \left(\varepsilon N a \dot{\beta} - \frac{a \dot{a}^2 \beta}{N} - \frac{a^2 \dot{a} \dot{\beta}}{N} \right)$$

up to a multiplicative constant. Let us point out that the constraint (115) can be obtained by variation of this lagrangian with respect to N .

The equations of motion obtained from the action (116) form a system which is equivalent to the system (114). This result validates the interpretation of our model as a minisuperspace model¹⁰.

The calculation of the Hamiltonian from the action (116) yields:

$$H = N \left[-\frac{\Pi_a \Pi_\beta}{a^2} + \beta \frac{\Pi_\beta}{a^3} + \beta a \right] =: N \mathcal{H} \quad (117)$$

where Π_a and Π_β are respectively the momenta associated to the variables a and β , and N plays the role of a Lagrange multiplier. Variation of the above hamiltonian with respect to N yields the minisuperspace version of the super-hamiltonian constraint (54), $\mathcal{H} \approx 0$, which is nothing but the constraint (115).

We can notice that, if we proceed through the Dirac quantization of our model using the variables (a, β) , the Hamiltonian (2.9) leads to factor-ordering problems¹¹. In this minisuperspace model, it is easy to circumvent such a difficulty. Let us introduce a new equivalent set of coordinates (x, y) and set

$$\begin{cases} x & := \beta a \\ y & := \frac{a^2}{2} \end{cases} \quad (118)$$

The action (116) is then given by

$$S = \int dt \left[\varepsilon N x - \frac{1}{N} \dot{x} \dot{y} \right] \quad (119)$$

The general solutions to the equations of motion are:

$$y = -\frac{\varepsilon t^2}{2} + \Sigma \quad (120)$$

$$x = ct \quad (121)$$

¹⁰In general, the simple substitution of an ansatz into the complete action of the theory gives an action whose equations of motion are not equal to the equations of motion of the full theory restricted to the ansatz. These procedures may not commute. This is usually the case when the constraints (55) are not identically zero. Therefore, we must be very careful with the procedure of obtaining the correct minisuperspace equations.

¹¹This is characteristic of the Wheeler-DeWitt equation (73) and it is one of the problems we have to face in the canonical quantization scheme [66, 67, 46].

which can be expressed as

$$y = -\varepsilon \frac{x^2}{2c^2} + \Sigma \quad (122)$$

where c and Σ are integration constants.

As it can be seen from equation (120), we may have the following possible classical solutions:

a) For $\varepsilon = 1$.

For $\Sigma > 0$, there is a singularity on $t = -\sqrt{\Sigma}$ when the universe is created, it expands till maximum size at $t = 0$, and then recollapse at $t = \sqrt{\Sigma}$.

If $\Sigma \leq 0$, there is no classical solution.

b) For $\varepsilon = -1$

If $\Sigma > 0$, the universe is flat at $t \rightarrow -\infty$, contracts to its minimum size at $t = 0$ and then expands to become flat again at $t \rightarrow \infty$. No singularities are present: it is an eternal universe.

If $\Sigma = 0$, it is just the flat spacetime in Milne coordinates.

If $\Sigma < 0$, we may have a universe that contracts from flat spacetime till a singularity or an expanding universe coming from a singularity and going to flat spacetime.

Thus, for $\varepsilon = -1$, there is the possibility of having eternal or singular universes, depending on the constant of integration Σ .

It is here where quantum cosmology enters. The idea is to apply the proposal of Ref. [62], which sustains that semi-classical wave functions are peaked on the correlations $p_i = \frac{\partial S(q_j, t)}{\partial q_i}$, and apply to this problem in order to see if one of the cosmological classical solutions (eternal or singular) of the problem we are now studying can be selected. We are taking this minisuperspace example because most of the qualitative features of minisuperspace quantum cosmology can be easily discussed due to the simple calculations it involves.

Let us then quantize the theory.

The Hamiltonian of the new action (119) is given by

$$H = -N(\Pi_x \Pi_y + \varepsilon x) \quad (123)$$

which yields the super-Hamiltonian constraint

$$\mathcal{H} := -(\Pi_x \Pi_y + \varepsilon x) \approx 0 \quad (124)$$

where $\Pi_x = -\frac{\dot{y}}{N}$ and $\Pi_y = -\frac{\dot{x}}{N}$.

The quantum version of equation (124) yields the minisuperspace Wheeler-DeWitt equation, which governs the dynamics of the quantum state $\psi(x, y)$. For our model this equation is given by:

$$\hat{\mathcal{H}} \left(x, -i \frac{\partial}{\partial x'}, -i \frac{\partial}{\partial y} \right) \psi(x, y) = 0 \quad (125)$$

where $\hat{\mathcal{H}}$ is the operator version of equation (124).

The explicit form of equation (125) is:

$$-\frac{\partial^2 \psi}{\partial x \partial y} + \varepsilon x \psi = 0 \quad (126)$$

A solution of (126) is given by

$$\psi(x, y) = \psi_0 \exp \left[\sqrt{-\varepsilon} \left(z_0 y - \frac{x^2}{2z_0} \right) \right] \quad (127)$$

where ψ_0 and z_0 are arbitrary complex constants.

Note that ψ is an eigenfunction of the momentum operator Π_y with eigenvalue $-i\sqrt{-\varepsilon}z_0$. In order for this eigenvalue to be real, z_0 must be real for $\varepsilon = 1$ or pure imaginary for $\varepsilon = -1$.

Thus, we may write equation (127) as:

$$\psi(x, y) = \psi_0 \exp \left[-i \left(cy - \varepsilon \frac{x^2}{2c} \right) \right] \quad (128)$$

where c is a real constant. This wave function is also a semi-classical wave function because the argument in the exponential of equation (128)

$$S \equiv -cy + \varepsilon \frac{x^2}{2c} \quad (129)$$

is the complete solution of the Hamiltonian-Jacobi equation of the model

$$\frac{\partial S}{\partial x} \frac{\partial S}{\partial y} + \varepsilon x = 0 \quad (130)$$

The general solution of equation (126) constructed from the particular solution (128) is:

$$\begin{aligned} \psi(x, y) &= \int dc G(c) \exp \left[-i \left(cy - \varepsilon \frac{x^2}{2c} \right) \right] \\ &= \int dc F(c) \exp \left[-i \left(cy - \varepsilon \frac{x^2}{2c} + \beta(c) \right) \right] \end{aligned} \quad (131)$$

where $G(c) = |G(c)|e^{-i\beta(c)} = F(c)e^{-i\beta(c)}$ is an arbitrary complex function.

We are interested in the wave function of the universe leading to a classical universe. In particular we want to know which of the possibilities of having an eternal or a singular classical universe is predicted. The occurrence of eternal or singular solutions depends on the sign of the constant ε , as discussed above. As it is impossible, in the above model, to obtain a classical eternal universe with $\varepsilon = 1$ we will, from now on, limit our discussion to the case $\varepsilon = -1$.

The idea that a quantum solution predicts a classical universe is meaningful, of course, only in the semi-classical limit which will be identified here with the behaviour of the wave function in the region where the scale factor is very large. Both quantities x^2 and y are proportional to a^2 (a is the scale factor) and so is the term $\left(cy + \frac{x^2}{2c} \right)$ in the phase of (131). Hence, when $a \rightarrow \infty$, this phase varies rather rapidly, enabling us to approximate $\psi(x, y)$, in the semi-classical limit, employing the stationary phase method. The stationary phase condition applied to equation (131) yields:

$$\frac{d\beta(c)}{dc} + y - \frac{x^2}{2c^2} = 0 \quad (132)$$

Suppose that equation (132) have N solutions, $c_n(x, y), n = 1, \dots, N$. In the semi-classical limit, $\psi(x, y)$ can be written as

$$\psi_{sc}(x, y) = \sum_{n=1}^N F(c_n(x, y)) \exp(-iS_n(x, y)) \quad (133)$$

with S_n defined by

$$S_n(x, y) := - \left[\beta(c_n(x, y)) + yc_n(x, y) + \frac{x^2}{2c_n(x, y)} \right] \quad (134)$$

It is easy to show that the $S_n(x, y)$ given by equation (134), are solutions of equation (130) with $\varepsilon = -1$:

$$\frac{\partial S}{\partial x} \frac{\partial S}{\partial y} - x = 0 \quad (135)$$

In the theory of the nonlinear partial differential equations of first order [68], it is called the general integral of equation (135). This type of equation admits two other sets of solutions: the complete and the singular integral. The complete integral of equation (135) is given by:

$$S_c(x, y) = -cy - \frac{x^2}{2c} \quad (136)$$

which is the solution (130) for $\varepsilon = -1$. There is no singular integral of equation (135).

The functions $S_c(x, y)$ given by (136) can be used to construct another set of semi-classical wave functions by the WKB approximation method. These WKB wave functions are approximate solutions in first order of \hbar of equation (125) (with $\varepsilon = -1$) and have the form:

$$\psi_W(x, y) = A(x, y) \exp[iS(x, y)] \quad (137)$$

For the particular case of equation (126), the functions $A(x, y)$ and $S(x, y)$ must satisfy, besides the Hamiltonian-Jacobi equation (135), the following equation

$$\frac{\partial A}{\partial x} \frac{\partial S}{\partial y} + \frac{\partial A}{\partial y} \frac{\partial S}{\partial x} + \frac{\partial^2 S}{\partial x \partial y} A = 0 \quad (138)$$

Using $S_c(x, y)$ given by (136) (which satisfies equation (135)), noting that $\frac{\partial^2 S_c}{\partial x \partial y} = 0$, and solving the above equation by the separation of variables method ($A(x, y) = X(x)Y(y)$), we obtain for the prefactor $A(x, y)$

$$A(x, y) = B \exp \left[W \left(-cy + \frac{c^2}{2c} \right) \right]$$

where B and W are real constants.

Thus, we obtain for the WKB wave function (137):

$$\psi_W(x, y) = B \exp \left[W \left(-cy + \frac{x^2}{2c} \right) - i \left(cy + \frac{x^2}{2c} \right) \right] \quad (139)$$

The semi-classical wave function $\psi_{sc}(x, y)$ given by equation (133) also satisfies the WKB equations (135) and (138). To see this, differentiate equation (132) with respect to x and y to obtain the following differential equation for $c(x, y)$

$$\frac{\partial c}{\partial x} = -\frac{x}{c^2} \frac{\partial c}{\partial y}$$

Using again the separation of variables method, we can obtain the general solution

$$c(x, y) = \pm \frac{\sqrt{x^2 + 2\gamma}}{\sqrt{2(y - \alpha)}} \quad (140)$$

where γ and α are constants of integration.

The functions $\beta(c)$ that yield these solutions are of the form

$$\beta(c) = -\alpha c + \frac{\gamma}{c}$$

Substituting into equation (134), the functions $S(x, y)$ are:

$$S_{\pm}(x, y) = \pm \sqrt{2(y - \alpha)} \sqrt{x^2 + 2\gamma} \quad (141)$$

It is easy to show that $S_{\pm}(x, y)$ satisfy equation (135), and that every smooth function $F(c)$ (the prefactor of ψ_{sc}) together with $S_{\pm}(x, y)$, satisfy equation (138) when $a \gg 1$. Also, $|\vec{\nabla} F| \ll 1$ for $a \gg 1$.

The wave functions ψ_{sc} and ψ_W given by equations (133) and (139), respectively, are the most general forms of WKB solutions in the form $\psi \sim e^{iS}$ where S is a solution of the Hamilton-Jacobi equation. (The prefactors should satisfy the relations $|\vec{\nabla} F| \sim |\vec{\nabla} A| \ll 1$).

As mentioned above, these wave functions have a peak on the correlations $p_i = \frac{\partial S}{\partial q_i}$, where the q^i are the minisuperspace variables and the p_i their canonical momenta. These correlations are in fact first integrals of the classical equations of motion. For the semi-classical wave functions given in equations (133) and (141) they yield:

$$\Pi_x = \frac{\partial S_{\pm}}{\partial x} = \pm \frac{x \sqrt{2(y - \alpha)}}{\sqrt{x^2 + 2\gamma}} \quad (142)$$

$$\Pi_y = \frac{\partial S_{\pm}}{\partial y} = \pm \frac{\sqrt{x^2 + 2\gamma}}{\sqrt{2(y - \alpha)}} \quad (143)$$

In the gauge $N = 1$, $\Pi_x = -\dot{y}$ and $\Pi_y = -\dot{x}$. Hence, the above equations yield:

$$\frac{dy}{dx} = \frac{\Pi_x}{\Pi_y} = \frac{\frac{\partial S_{\pm}}{\partial x}}{\frac{\partial S_{\pm}}{\partial y}}$$

whose general solution is

$$y = \frac{x^2}{2c^2} + \frac{\gamma}{c^2} + \alpha \quad (144)$$

where c is a real integration constant.

Looking to the above equation we see that $\Sigma = \frac{\gamma}{c^2} + \alpha$. Therefore, if we know the exact form of $S_{\pm}(x, y)$ (or $\beta(c)$), it is possible to make exact predictions about the singular nature of the classical solutions. If both α and γ are positive, negative or null, the wave functions predict eternal, singular or Minkowski universes, respectively.

Note that, in these cases, the knowledge of the function $S(x, y)$ is enough to make exact predictions.

The other type of semi-classical wave functions given in equation (139) are peaked on

$$\Pi_x = \frac{\partial S_c}{\partial x} = -\frac{x}{c} \quad (145)$$

$$\Pi_y = \frac{\partial S_c}{\partial y} = -c \quad (146)$$

where S_c is given by (136), which yields the following first integral of the classical equations of motion:

$$\frac{dy}{dx} = \frac{x}{c^2} \Rightarrow y = \frac{x^2}{2c^2} + \Sigma \quad (147)$$

The constant Σ comes again as an integration constant. Therefore, its sign is still unknown. The knowledge of the function $S(x, y)$ is not enough to make predictions.

However, in the semi-classical approximation, we can use the prefactor to define a measure over the ensemble of classical trajectories in the minisuperspace of the problem, around which the wave function is peaked. This is done in the following way: the WKB solutions are of the form $\psi = A(q)e^{iS(q)}$. We can construct the minisuperspace vector (a vector which is defined in minisuperspace whose indices are raised and lowered by the Wheeler-DeWitt metric)

$$j_i = A^2(q^j)\nabla_i S(q^j) \quad (148)$$

where the ∇_i is the covariant derivative with respect to the Wheeler-DeWitt metric. Using the Wheeler-DeWitt equation in the WKB approximation, it can easily be shown that this vector has null covariant divergence [82]. Then we can construct a conserved measure on minisuperspace

$$P = \int dP = \int j_i d\sigma^i \quad (149)$$

if $d\sigma^i$ is the “area element” of a suitably chosen hypersurface of minisuperspace, so that all the trajectories of the ensemble of classical trajectories cross it only once.

Once we have defined this measure, we can calculate what is the relative measure of some of the classical solutions with respect to the others. If this relative measure is very close to one, we will say that this classical solution is predicted. If it is very close to zero, it is excluded.

Let us return to our example. We define:

$$\eta = y - \frac{x^2}{2c^2} \quad (150)$$

$$\xi = -y - \frac{x^2}{2c^2} \quad (151)$$

We then obtain for j_i given in equation (148)

$$j_i = \exp(2cW\eta)\nabla_i(c\xi) \quad (152)$$

In the plane (ξ, η) , the surfaces $\xi = \text{const.}$ (ξ is essentially S_c given in (136)) are orthogonal to the classical trajectories (147) (by virtue of equations (145) and (146)) which have $\eta = \Sigma = \text{const.}$ Thus, the vector \vec{j} points largely in the ξ direction. Choosing the surfaces σ in (169) to be the surfaces of constant ξ , it will be guaranteed that \vec{j} crosses them only once. This choice yields for (149) the following equation:

$$dP = \vec{j} \cdot d\vec{\sigma} \approx \exp(2cW\eta)d\eta$$

The conditional probability of having $\eta \approx \Sigma > 0$ will be given by:

$$P(\eta \approx \Sigma > 0 | -\infty < \eta < \infty) = \frac{\int_0^\infty \exp(2cW\eta)d\eta}{\int_{-\infty}^\infty \exp(2cW\eta)d\eta} \quad (153)$$

If $cW > 0$ then $P = 1$ and if $cW < 0$, $P = 0$. In each case the condition $|\nabla A| \ll 1$ as $a \gg 1$ is satisfied. Thus, we can make a definite prediction about the sign of Σ if we know the sign of cW , which is given by the wave function (139).

We have seen that both wave functions (133) (with (141)) and (139) select one and only one classical solution depending on the values of the constants γ and α in the first case, and c and W in the second case¹². To know these constants, we need boundary conditions on the Wheeler-DeWitt equation in order to select one and only one solution with their specifics (γ, α) or (c, W). The reader may object by saying that we have only displaced the problem: why not impose boundary conditions on Einstein's equations directly? This is because our experience in quantum mechanics shows that it is common to have natural boundary conditions on the wave function (usually coming from impositions of regularity on the boundaries of the space where its arguments are defined), which sometimes are sufficient to select a unique wave function. Also, if there is some deep principle which explains the puzzles in classical cosmology, it must be sought in the more fundamental quantum theory. We will return to these problems in the next section.

Another important problem with the program suggested in Ref. [62], as we have mentioned before, is that *it is not true* that WKB wave functions have their Wigner functions peaked on the correlations $p - \frac{\partial S(x,t)}{\partial x}$. The approximation method used to arrive at equation (106) is not correct. In fact, in Ref. [61] it is shown that using the uniform approximation method for one dimensional systems, the Wigner function of semi-classical wave functions is proportional to Airy functions, which may have a lot of peaks. As discussed in Ref. [69], these Airy functions are much more close to the exact Wigner functions calculated in some simple examples than the Dirac-delta function (106). Therefore, if semi-classical Wigner functions may have many peaks, no prediction can be made out of them. Fortunately, this is not the case of the minisuperspace example we have discussed in this section but in general we will have to face this problem. This point will also be discussed in the next section.

¹²In the first case, it is possible to make predictions only with the knowledge of the hamilton-jacobi function (141); in the second case, the pre-factor (138) together with a definition of measure on minisuperspace are also necessary.

5 Boundary conditions and decoherence

5.1) The no-boundary boundary condition

As was explained in the last section, in order to quantum cosmology be a theory of initial conditions of the universe, boundary conditions to the Wheeler-DeWitt equation are needed in order to select a unique solution which will be called *the* wave function of the universe. The minisuperspace example we have studied above suggested this need¹³. There are many proposals of boundary conditions to the Wheeler-DeWitt equation [71, 72], the most popular being the tunnelling boundary condition [73, 74] and the no-boundary boundary condition [75, 76]. Only the second one will be discussed here.

The no-boundary boundary condition is defined in the context of what is called euclidean quantum gravity [77]. It makes use of a path integral formulation of quantum gravity. In ordinary quantum mechanics, a solution of the Schrödinger equation can be written as a path integral in the following way:

$$\begin{aligned}\Psi(x, t) &= \int dx_0 K(x, t; x_0, t_0) \Psi(x_0, t_0) \\ &= \int \int D_F[x(\tau)] \exp\left\{\frac{i}{\hbar} S([x(\tau)] : x, t; x_0, t_0)\right\} \Psi(x_0, t_0) dx_0\end{aligned}\quad (154)$$

where $K(x, t; x_0, t_0) = \langle x|U(t, t_0)|x_0 \rangle$ is the propagator, $\Psi(x_0, t_0)$ is the initial wave function (which reflects the way the system has been prepared), and the integration $\int D_F[x(\tau)]$ is over all possible paths between (x_0, t_0) and (x, t)

The ground state wave function can be obtained in the following way: choose $t = 0$, $x_0 = 0$, write $t' = t_0$, substitute in the propagator, and insert in it the identity operator written in terms of a complete set of energy eigenfunctions:

$$\begin{aligned}K(x, 0; 0, t') &= \langle x|U(0, t')|0 \rangle = \sum_n \langle x|U(0, t')|\varphi_n \rangle \langle \varphi_n|0 \rangle \\ &= \sum_n \varphi_n(x) \varphi_n(0) \exp(iE_n t') \\ &= \int D_F x(\tau) \exp\left\{\frac{i}{\hbar} S([x(\tau)] : x, 0; 0, t')\right\}\end{aligned}\quad (155)$$

If we make a Wick rotation $t' = -i\tau'$ and take the limite $\tau \rightarrow -\infty$, the unique term in the sum which survives is the one with the lowest energy. Thus, we can write the ground state wave function as:

$$\varphi_0(x) = N \int D_F x(\tau) \exp\left\{-\frac{1}{\hbar} I([x(\tau)] : x, 0; 0, \tau' \rightarrow -\infty)\right\}\quad (156)$$

where $I(\tau) \equiv -S(t = -i\tau)$ is the Euclidean action, and N is a normalization constant. When the action is the kinetic energy T minus the potential energy V , $S = T - V$, the

¹³However, Ref. [70] shows a minisuperspace example where a natural measure can be defined in the space of solutions of the Wheeler-DeWitt equation. Using this measure, they show that inflation is a common feature of these solutions. This is an example of prediction without boundary conditions.

euclidean action turns out to be $I = T + V$, the total energy and, as a consequence, positive definite. Consequently, the path integral in equation (156) converges.

In quantum field theory, we have an analogous equation for the ground state wave functional:

$$\Psi_0[\phi(x), 0] = N \int D_F[\phi(x)] \exp\left\{-\frac{1}{\hbar} I([\phi(x)])\right\} \quad (157)$$

where the path integral is over all Euclidean field configurations $\phi(x, \tau)$ to the past of $\tau = 0$ which match with $\phi(x)$ in the hypersurface $\tau = 0$. As in ordinary quantum mechanics, the Euclidean actions of usual field theories are positive definite.

In general relativity, the notion of gravitational energy has a meaning only in spacetimes with timelike Killing vectors. Hence, we cannot use the notion of energy in order to define the ground state. The proposal of Hartle and Hawking [75] is to extend the Euclidean functional integral definition of ground state in ordinary quantum mechanics and quantum field theory to the domain of general relativity:

$$\Psi_0[h_{ij}(x), \phi(x), B] = N \int D_F[h_{ij}(x)] D_F[\phi(x)] \exp\left\{-\frac{1}{\hbar} I([g_{\mu\nu}(x), \phi(x)])\right\} \quad (158)$$

where the integration is over all Euclidean four-geometries and Euclidean field configurations which match with $h_{ij}(x)$ and $\phi(x)$ in the three-boundary B .

As was discussed in section 3, if the four-geometry is spatially open, the action of general relativity has surface terms. For cosmological geometries, whose metric components do not go to zero at spatial infinity, these surface terms are usually infinity. The Euclidean action diverges like the spatial volume of the open universe. Therefore, in the path integral (158), the contributions of spatially open geometries seems to be zero (or infinity but then the wave function will be ill defined), and so we will only consider spatially closed universes. This is the first feature of the wave function (158); it seems to predict that the universe is spatially closed (evidently, this is not a rigorous statement). It is in accordance with the conjecture which says that this wave function is a kind of ground state wave function: a closed universe is the closest one to a vacuum state because its energy and all charges are zero.

The path integral in equation (158) will be over spatially closed Euclidean four-geometries. We can parametrize the three-geometries h_{ij} and field configurations ϕ of this four-geometry by a parameter λ in such a way that $h_{ij}(x, \lambda = 1) = h_{ij}(x)$; $\phi(x, \lambda = 1) = \phi(x)$, the arguments of the wave function (158). In order to obtain a unique Ψ_0 out of equation (158), we must specify what are the values of geometry and matter fields at some other value of λ , at some other boundary. The proposal of Hartle and Hawking is that there is no other boundary; the integration in (158) must be over Euclidean four-geometries which are compact in space *and* time. The sole boundary is the one which appears in the argument of the wave function. The universe in the Euclidean regime has no boundary in space or time. That is why it is called the no-boundary proposal. The foliation of this compact Euclidean four-geometry will necessarily reach the ‘south pole’ of this geometry at some λ , say $\lambda = 0$. At this point, $h^{1/2} = 0$ but the geometry is perfectly regular there. Thus, the boundary conditions at $\lambda = 0$ are $h^{1/2}(\lambda = 0) = 0$, and conditions of regularity on the matter fields and on the derivatives of the geometry and matter fields at this point.

There is a problem, however, with equation (158). The euclidean action of general relativity *is not* positive definite. In fact, as we have seen in section 3, the kinetic term of the general relativity lagrangean has a metric which is not positive definite: the Wheeler-DeWitt metric. As we have discussed there, the negative sign of the Wheeler-DeWitt metric corresponds to the determinant of the spacelike metric h_{ij} . It means, as can easily be checked, that if we have a metric for which the Euclidean action is positive definite, a conformally related metric $\tilde{g}_{\alpha\beta} = \Omega^2 g_{\alpha\beta}$ may have a negative action depending on Ω being a rapidly varying function.

To solve this problem, we have to distort the contour of integration in equation (158) to complex metrics and find a contour which gives a finite Ψ_0 . However, this contour is not unique and we may have different finite no-boundary wave functions for different contours [78]. This is a serious problem of the no-boundary proposal; it may not give a unique wave function of the universe. This motivated some authors to add some new restrictions to the no-boundary proposal [79]. We will not enter on these details here.

As general relativity is a theory with constraints, the path integral in equation (158) must be made with care in order to not integrate over spurious degrees of freedom, like in any other gauge field theory. For details on this point see Refs. [80, 43, 44].

In the semi-classical approximation, the dominant contributions to the path integral (158) will come from the four-geometries and matter fields which are solutions of the Euclidean equations of motions because they minimize the Euclidean action:

$$\Psi_0^{sc}[h_{ij}(x), \phi(x), B] \propto \sum_n \exp\left\{-\frac{1}{\hbar} I_n^{cl}([g_{\mu\nu}(x), \phi(x)])\right\} \quad (159)$$

where the sum is over Euclidean classical solutions.

Let us impose the no-boundary boundary condition to the solutions of the minisuperspace model we have presented in the last section to see how it works. The no-boundary semi-classical wave function can be calculated by evaluating the Euclidean action that comes from (119) (with $\varepsilon = -1$)

$$I = - \int d\tau \frac{1}{2} \left(\frac{\dot{x}\dot{y}}{N} - Nx \right)$$

in the classical solutions of the Euclidean field equations

$$\left(\frac{\dot{y}}{N}\right)' + N = 0$$

$$\left(\frac{\dot{x}}{N}\right)' = 0$$

$$\frac{\dot{x}\dot{y}}{N^2} + x = 0$$

At the boundary B , where the wave function is evaluated, we have $y(1) = \tilde{y}$ and $x(1) = \tilde{x}$. At the ‘south pole’, labelled by $\tau = 0$, the no-boundary proposal says that $h^{1/2} = 0$. This implies that $a(0) = y(0) = 0$. Also, it requires regularity of the fields there: $|\phi(0)| < \infty$ which implies that $x(0) = a(0)[1 - \sigma\phi^2(0)] = 0$. With these boundary conditions, the solutions of the Euclidean equations of motion are: $y = \tilde{y}t^2$, $x = \tilde{x}t$ and

$N = \pm i\sqrt{2\tilde{y}}$. These solutions can be substituted into the action yielding $I_{\pm} = \pm i\tilde{x}\sqrt{2\tilde{y}}$. The semi-classical no-boundary solution is given by:

$$\psi_{NB} \propto \exp(-I_+) + \exp(-I_-) \propto \cos(x\sqrt{2y}) \quad (160)$$

This is a wave function of the type ψ_{sc} with S given by equation (141) with $\alpha = \gamma = 0$. Therefore, by equation (144), the no-boundary wave-function predicts Minkowski space-time.

This result is in accordance with the conjecture that the no-boundary wave function is a kind of vacuum wave function in the sense that it represents a state with maximal symmetry. In fact, Minkowski spacetime is the solution with maximum number of Killing vectors in our minisuperspace model.

For the $\varepsilon = +1$ case the solution is

$$\psi_{NB} = \exp(x\sqrt{2y}) \quad (161)$$

It is not an oscillatory wave function. It can be viewed as generating solutions of equation (120) with $\Sigma = 0$, which are not allowed for $\varepsilon = +1$. In general, oscillatory wave functions are related to classical solutions while non-oscillatory wave functions are related to classical forbidden regions, like tunneling effects, exactly like in ordinary quantum mechanics.

The no-boundary boundary condition, as other boundary conditions, has been applied to many different minisuperspace examples as attempts to answer old cosmological questions, as described in the introduction.

For minisuperspaces of homogeneous but anisotropic universes, like Bianchi I and Bianchi IX models, it was shown that the no-boundary wave function has a peak where the minisuperspace coordinates describe isotropic cosmological solutions [81].

In the case of a homogeneous minimally coupled scalar field, some authors have studied if the no-boundary and tunneling solutions yield the initial conditions for having inflation. As explained in the lectures of Branderberger, inflation can explain the observed isotropy of the universe. However, not all initial conditions give rise to inflation; depending on the initial value of the scalar field, the universe may recollapse too early or we may not have enough inflation to isotropize it. In Ref. [82] this issue is discussed comprehensively, and it is shown that the tunneling solution yields good inflation while the no-boundary does not (although some controversy on this result exists in the literature).

The problem of structure formation can also be discussed in the light of quantum cosmology [82, 83]. Small inhomogeneous perturbations are added to the homogeneous and isotropic metric and scalar field, enlarging the minisuperspace model:

$$\begin{aligned} h_{ij}(x, t) &= a^2(t)(\Omega_{ij} + \epsilon_{ij}(x, t)) \\ \Phi(x, t) &= \phi(t) + \delta\phi(x, t) \\ N(x, t) &= N_0(t) + \delta N(x, t) \\ \delta N^i(x, t) & \end{aligned} \quad (162)$$

where Ω_{ij} is the metric on the unit three sphere.

These inhomogeneous perturbations are expanded in three-sphere harmonics Q_{lm}^n , as for example the scalar field:

$$\delta\phi(x, t) = \sum_{nlm} f_{nlm}(t) Q_{lm}^n(x) \quad (163)$$

The new minisuperspace action is:

$$S(g_{\mu\nu}, \Phi) = S_0(N_0, a, \phi) + S_2(N_0, a, \phi, \delta N, N^i, f_{nlm}^\alpha) \quad (164)$$

where f_{nlm}^α represents all expansion coefficients of the perturbations.

We can proceed with the quantization of this enlarged minisuperspace in the usual way (see details in the Refs. [82, 83]). In the semi-classical approximation, we can obtain a Schrödinger equation for a wave function describing the quantum evolution of the perturbations in a fixed curved spacetime, following the same steps we have taken in order to obtain equation (88) in section 3. This wave function can be obtained from the original solution of the Wheeler-DeWitt equation, as in equations (82), (86) and (87). If a unique solution of the Wheeler-DeWitt equation is selected by some boundary condition (no-boundary, tunneling, etc), the quantum state of primordial perturbations can be known and its evolution uniquely determined. After, it can be compared with observations as explained in Refs. [84, 85].

5.2) Decoherence

Let us now return to the problem of finding peaks in the Wigner function, as discussed in the last section. As Berry have shown, Wigner functions of semi-classical wave functions do not have, in general, a unique peak. Also, it may sometimes happen that the semi-classical wave function be a superposition of many WKB wave functions, like in equation (160), and the corresponding Wigner function present quantum interference (which is not the case of the particular precedent case for the questions we were trying to answer).

Usually, quantum interference can be erased by the phenomenon of decoherence. It happens when the degrees of freedom under study are in interaction with a macroscopic environment. When all the irrelevant degrees of freedom of the environment are traced out, we obtain a reduced density matrix, like in equation (37) of section 2. As we have pointed out in that section, the term responsible for quantum interference vanishes for almost every macroscopic system [25, 26, 27, 28, 29].

The natural idea is to apply the concept of decoherence to quantum cosmology. But what is the ‘environment’ of the universe? Our observations of the universe are coarse. Therefore, we could think the ‘environment’ of the universe as composed of fine degrees of freedom which are not observed. In Ref. [86], a minisuperspace example of this idea has been developed. It was also shown in this paper that the Wigner function of the reduced density matrix obtained by tracing out the unobserved degrees of freedom has only one peak. They took a very simple minisuperspace model where the unique degree of freedom was the scale factor. The theory is general relativity with a cosmological constant. The Wheeler-DeWitt equation is:

$$\left[4\frac{d^2}{dq^2} - 1 + \Lambda q\right]\Psi(q) = 0 \quad (165)$$

where $q = a^2$, a being the scale factor, and Λ is the cosmological constant. The no-boundary wave function is $\Psi_{NB} = -i\text{Ai}[\frac{1-\Lambda q}{(2\Lambda)^{2/3}}$]. The semi-classical Wigner function of the no-boundary wave function is proportional to a sinus function and has many peaks. After, they introduced inhomogeneous perturbations, like in equation (162). They calculated the modified density matrix, integrated over the inhomogeneous degrees of freedom to obtain the reduced density matrix, and calculated the semi-classical Wigner function of this reduced density matrix. The result was that, depending on the degrees of freedom that were traced out, the new Wigner function has only one peak over the classical trajectories.

Therefore, decoherence is not only responsible for the erasing of quantum interference but also for the construction of classical correlations between phase space variables.

There are many questions still to be answered. The calculations of Ref. [86] indicate that the tracing out of unobserved degrees of freedom is not arbitrary: if we trace out too many degrees of freedom, we end with no peak; if we trace out only a few, we end with a lot of peaks. Until now, there is no physical reason that can guide us with the good choice. This problem leads to another one: what is a good peak? The Dirac delta peaks of Ref. [62] are undoubtedly good peaks but they are not correct, in general, as we have seen. It is unavoidable to deal with finite peaks. Therefore, the question of how big must be a peak in order to yield exact predictions is pertinent, and without a general and satisfying answer. This motivates us to the next section, where we will present some other interpretations of quantum mechanics which may be more appropriate to quantum cosmology.

6 The ontological and consistent histories interpretations of quantum mechanics

What we call the consistent histories interpretation is the one developed by Griffiths, Omnès, Hartle and Gell-Mann [87, 21, 88, 89], and the ontological interpretation is the one developed basically by David Bohm [30, 31].

6.1) The consistent histories interpretation

The consistent histories interpretation is an improvement of the idea behind the many-worlds interpretation. Quantum mechanics is not viewed as a theory of many worlds but as a theory of many histories. It was developed by Griffiths and Omnès in order to get a consistent interpretation of quantum mechanics without the problems mentioned in section 2.

The first basic assumption of this scheme is that, according to Omnès [21], ‘every physical system, whether an atom or a star, is assumed to be described by a universal kind of mechanics, which is quantum mechanics’. There are two immediate important consequences of this assumption: first that the theory deals with individual systems (there is no sense in dealing with an ensemble of planets Mars in order to study this planet), and second that classical mechanics must be derived from quantum mechanics in the situations where it is a good approximation. Here, classical mechanics means not

only classical dynamics (Newton's laws, in the non-relativistic case) but also classical logic (common sense), determinism, and everything characteristic of the classical world. Therefore, the classical world must be derived from the quantum world.

Evidently, this kind of interpretation is better suited to quantum cosmology than the Copenhagen one. That is why Hartle and Gell-Mann have developed an analogous framework in order to apply it to quantum cosmology.

In the history interpretation, probabilities¹⁴ are not assigned to events as in usual quantum mechanics but to whole histories. However, as we know, we cannot assign probabilities to every history in quantum mechanics. The interference figure obtained from the two slit experiment is an evidence of this fact. Hence, we must establish what are the conditions on families of histories in order to be possible to assign probabilities to all members of such families. Once we obtain these conditions, we will have the possibility of saying, for instance, that a history of the universe with inflation is more probable than another one without inflation, without mentioning observers or measurements. Let us give more details on how this interpretation works.

A history of an isolated physical system is a succession of properties of this system occurring at different times. An example of a property of a system is the sentence 'the eigenvalue of the observable \hat{B} is in the set D '. To each property is associated a projector operator. In the above example, it would be the projector P onto the subspace of the Hilbert space containing all eigenvectors with eigenvalues in the set D . Another way to say the above property is 'the value of P is 1'.

The probability of a property, designed by its projector P , must satisfy the following conditions:

$$0 \leq p(P) \leq 1 \tag{166}$$

$$p(I) = 1 \tag{167}$$

$$p(P + P') = p(P) + p(P') \tag{168}$$

where P and P' are projectors into disjoint sets D and D' .

There is a theorem due to Gleason [90], which shows that there exists a trace-class (with unit trace) positive operator ρ (the density operator), where a $p(P)$ satisfying the above conditions can be written as (compare with equation (27)):

$$p(P) = \text{Tr}(\rho P) \tag{169}$$

The probability of a history can also be obtained from some logical conditions (for details, see Ref. [21]). The unique¹⁵ probability is given by:

$$p = \text{Tr}\{P_n(t_n)\dots P_k(t_k)\dots P_1(t_1)\rho P_1(t_1)\dots P_k(t_k)\dots P_n(t_n)\} \tag{170}$$

¹⁴Here, probability has only a formal meaning, a mathematical object which must satisfy some mathematical requirements, as will see later on. Its connection with the relative frequencies of measurement data is something to be established when a theory of measurements is formulated. It is argued in Ref. [21] that there are some probabilities which cannot be tested by measurements while there are others which may have an empirical sense.

¹⁵The uniqueness is only proved for histories with two instants of time or histories where the projectors refers either to position or momentum operators.

where ρ is the density matrix of the initial state of the system. One of the projectors $P_n(t_n)$ can be omitted due to the cyclic property of the trace and the fact that $P_n(t_n)$ is a projector. Note that for $n = 1$ this probability reduces to equation (169). Also, if ρ represents a pure state, $\rho = |\Psi\rangle\langle\Psi|$, this probability reduces to the reasonable equation:

$$p = |P_n(t_n)\dots P_k(t_k)\dots P_1(t_1)|\Psi\rangle|^2 \quad (171)$$

If we have more than one history, constituting what will be called a family of histories, then the additivity condition on probabilities must be checked. Let us make some definitions.

Two histories h and h' are said to be disjoint if $P'_k(t_k)P_k(t_k) = 0$ for some k . The union of two histories is defined if the histories have $P'_i(t_i) = P_i(t_i)$ for all i except for one $i = k$ where $P'_k(t_k)P_k(t_k) = 0$ (they must be disjoint). The union is the history given by the sequence $\{P_1(t_1)\dots P_i(t_i)\dots P_k(t_k) + P'_k(t_k)\dots P_n(t_n)\}$

A consistent family of histories is one where each probability of each possible union of two disjoint histories is the sum of the probabilities of each disjoint history:

$$p(h + h') = p(h) + p(h') \quad (172)$$

Therefore, in a consistent family of histories, a probability can be assigned to each history of the family. Equation (172) implies some consistency conditions. Let us examine a simple example. Take a family constituted of two histories and two instants of time. The history h is $\{P_1(t_1), P_2(t_2)\}$ and h' is $\{P'_1(t_1), P_2(t_2)\}$, with $P_1(t_1)P'_1(t_1) = 0$. The initial state is given by the density matrix ρ . Then we have:

$$\begin{aligned} p(h + h') &= \text{Tr}\{P_2(t_2)(P_1(t_1) + P'_1(t_1))\rho(P_1(t_1) + P'_1(t_1))\} \\ &= p(h) + p(h') + \\ &+ \text{Tr}\{P_2(t_2)P'_1(t_1)\rho P_1(t_1)\} + \text{Tr}\{P_2(t_2)P_1(t_1)\rho P'_1(t_1)\} \end{aligned} \quad (173)$$

Hence, probabilities can be assigned to this family of histories if:

$$\text{Tr}\{P_2(t_2)P'_1(t_1)\rho P_1(t_1)\} + \text{Tr}\{P_2(t_2)P_1(t_1)\rho P'_1(t_1)\} = 0 \quad (174)$$

Using that the projectors are hermitean operators, equation (174) is equivalent to:

$$\text{ReTr}\{P_2(t_2)P_1(t_1)\rho P'_1(t_1)\} = 0 \quad (175)$$

This is the consistency condition for this family of histories.

To illustrate the meaning of equation (175), let us apply it to a concrete example. Consider a spin $\frac{1}{2}$ particle. The initial state is given by the property $\sigma.n_0 = +1$ where n_0 is some unit length vector and σ are the Pauli matrices. At times t_1 and t_2 , the possible properties are given by $\sigma.n_j = \pm 1$, $j = 1$ or 2 , n_1 and n_2 being two others unit vectors. It can be shown that the consistency conditions derived from this family of histories imply that:

$$(n_0 \wedge n_1).(n_1 \wedge n_2) = 0 \quad (176)$$

where \wedge is the vector product.

i) Take $n_0 = x$, $n_1 = y$ and $n_2 = z$. In this case, condition (176) is satisfied. A history h is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$, $S_y = \frac{1}{2}\hbar$ at $t = t_1$ and $S_z = \frac{1}{2}\hbar$ at $t = t_2$. A history h' is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$, $S_y = -\frac{1}{2}\hbar$ at $t = t_1$ and $S_z = \frac{1}{2}\hbar$ at $t = t_2$. The union $h + h'$ is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$ and $S_z = \frac{1}{2}\hbar$ at $t = t_2$. The intermediate step at time t_1 is omitted because the sum of the projectors associated with this step in h and h' is the identity operator. In other words, the statement that the y component of the spin of the particle at time t_1 is either $\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$ is a trivial statement which can be omitted.

It can be easily verified that the probability of $h + h'$ is the sum of the probabilities of h and h' . Indeed:

$$\begin{aligned} p(h + h') &= p(h) + p(h') \\ \frac{1}{2} &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \end{aligned}$$

ii) Take $n_0 = x$, $n_1 = y$ and $n_2 = x$. In this case, condition (176) is *not* satisfied. A history h is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$, $S_y = \frac{1}{2}\hbar$ at $t = t_1$ and $S_x = \frac{1}{2}\hbar$ at $t = t_2$. A history h' is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$, $S_y = -\frac{1}{2}\hbar$ at $t = t_1$ and $S_x = \frac{1}{2}\hbar$ at $t = t_2$. The union $h + h'$ is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$ and $S_x = \frac{1}{2}\hbar$ at $t = t_2$. The intermediate step at time t_1 is omitted for the same reason as before.

It can be easily verified that the probability of $h + h'$ is *not* the sum of the probabilities of h and h' .

$$\begin{aligned} p(h + h') &\neq p(h) + p(h') \\ 1 &\neq \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \end{aligned}$$

iii) Take $n_0 = x$, $n_1 = x$ and $n_2 = z$. In this case, condition (176) is satisfied. A history h is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$, $S_x = \frac{1}{2}\hbar$ at $t = t_1$ and $S_z = \frac{1}{2}\hbar$ at $t = t_2$. A history h' is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$, $S_x = -\frac{1}{2}\hbar$ at $t = t_1$ and $S_z = \frac{1}{2}\hbar$ at $t = t_2$. The union $h + h'$ is: the particle has spin $S_x = \frac{1}{2}\hbar$ at $t = t_0$ and $S_z = \frac{1}{2}\hbar$ at $t = t_2$.

The probability of $h + h'$ is the sum of the probabilities of h and h' .

$$\begin{aligned} p(h + h') &\neq p(h) + p(h') \\ \frac{1}{2} &= 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} \end{aligned}$$

For more complicate families of histories, the necessary and sufficient consistency conditions are more involved [21]. That is why Hartle and Gell-Mann [88] prefer to use a simpler sufficient, but not necessary, condition. They defined the 'decoherence functional' as:

$$D(\{P_{\alpha'}\}\{P_{\alpha}\}) = \text{Tr}\{P_{\alpha'_n}^n(t_n)\dots P_{\alpha'_1}^1(t_1)\rho P_{\alpha_1}^1(t_1)\dots P_{\alpha_n}^n(t_n)\} \quad (177)$$

(the indices α_n are to emphasize that we may have many projectors at each instant t_n).

Their sufficient condition is:

$$D(\{P_{\alpha'}\}\{P_{\alpha}\}) = 0 ; \alpha_{k'} \neq \alpha_k \quad (178)$$

This implies that the decoherence functional can be written as:

$$D(\{P_{\alpha'}\}\{P_{\alpha}\}) = \delta_{\alpha'_1\alpha_1}\dots\delta_{\alpha'_n\alpha_n}p(h = \{P_{\alpha}\}) \quad (179)$$

for each history $h = \{P_{\alpha}\}$ of the given family of histories.

Families of fine grained histories (for instance, precise values of the position operator at every instant of time) are not consistent. Usually we have to deal with coarse grained histories (for instance, values of the position operator belonging to some set of values at some instants of time). These coarse grained histories may satisfy, at least approximately, equation (178) (recall that our observations in cosmology are very coarse grained). In this case, we may assign probabilities to them. There must exist some families of coarse grained histories which satisfy equation (178) with no finer-grained family which satisfies it. These families are called maximal sets. The time evolution contained in some histories belonging to consistent families may be approximately equal to the time evolution obtained from the classical equations of motion. These are quasi-classical histories. They involve quasi-classical projectors associated with collective observables (e.g., the center of mass position of a collection of atoms).

In quantum cosmology, the goal would be to find collective observables (related with concrete observations), and their connections with fundamental quantum gravity operators, identify consistent family of histories, impose as initial condition some solution of the Wheeler-DeWitt equation obtained from some suitable boundary conditions as described in the precedent section, and finally calculate probabilities of histories. This is subject of intense research nowadays [91, 92, 93, 94, 95].

It should be emphasized that there are some important differences in the formulations of the history interpretation. The first we have already mentioned: while Hartle and Gell-Mann work with the simpler sufficient conditions (178), Omnès works with more complicate sufficient and necessary conditions. It means that Hartle and Gell-Mann loose some consistent families of histories in the name of simplicity.

Second, in the original formulation of Griffiths, the initial and final states in the histories are fixed while in the Omnès formulation the final state is open. This implies that the Griffiths formulation is time reversal while the Omnès one is not. Omnès argues that with his definition of probability, he can define a consistent logic in consistent families of histories while with Griffiths definition this is not possible. This is undoubtedly a very persuasive argument.

Let us now introduce another interpretation, the ontological interpretation of quantum mechanics.

6.2) The ontological interpretation

The ontological interpretation of quantum mechanics works as follows: take the Schrödinger equation in the coordinate representation with the hamiltonian $H = \frac{P^2}{2m} + V(x)$

$$i\hbar\frac{d\Psi(x,t)}{dt} = [-\frac{\hbar^2}{2m}\nabla^2 + V(x)]\Psi(x,t) \quad (180)$$

Write $\Psi = R \exp(iS/\hbar)$ and substitute in (180). We obtain the following equations:

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} = 0 \quad (181)$$

$$\frac{\partial R^2}{\partial t} + \nabla \cdot (R^2 \frac{\nabla S}{m}) = 0 \quad (182)$$

The ontological interpretation, based on these two equations, is the following [30]:

- i) The quantum particles follow trajectories $x(t)$, *independent of observations*.
- ii) The particles are never separated from a quantum field Ψ which acts on them, which satisfies the Schrödinger equation (180).
- iii) The momentum of the particle is $p = \nabla S$.
- iv) Equation (181) is a Hamilton-Jacobi type equation for a particle submitted to an external potential which is the classical potential plus a new quantum potential

$$Q \equiv -\frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} \quad (183)$$

Hence, the particle trajectory $x(t)$ satisfies the equation of motion

$$m \frac{d^2 x}{dt^2} = -\nabla V - \nabla Q \quad (184)$$

v) In a statistical ensemble of particles in the same quantum field Ψ , the probability density is $P = R^2$. Equation (182) guarantees its conservation on time.

Let us make some comments:

a) Even in the regions where Ψ is very small, the quantum potential can be very high, as we can see from equation (183). It depends only on the form of Ψ , not on its absolute value. This fact brings home the non-local character of the quantum potential. This is very important because the Bell's inequalities together with Aspect's experiments show that, in general, a quantum theory must be either non-local or non-ontological. As Bohm's interpretation is ontological, it must be non-local, as it is. The quantum potential is responsible for the quantum effects. For instance, in the two-slit experiment, the gradient of the quantum potential, the quantum force, is infinite exactly on the points of destructive interference; particles cannot be there.

b) An image proposed by Bohm and Hiley is that the wave function Ψ acts like a radio wave emitted to an automatic pilot in a ship and guide it. It has not the energy to pull the ship but it gives the information for its engine, in our case the quantum potential, to do so.

c) It is not always true that we can write the probability density of an statistical ensemble of quantum particles as $P = R^2$. The function R is more important to construct the quantum potential than to assign probabilities. Probabilities are not fundamental in this interpretation.

d) The classical limit is very simple: we have only to find the conditions for having $Q = 0$.

e) As we have discussed in section 2, in a measurement the wave function is a superposition of non-overlapping wave functions, as we can see from equation (34) and equation

(35). The particle will enter in only one region and it will be influenced by the quantum potential obtained from the non-zero wave function of this region only.

For quantum fields, we can apply a similar reasoning. As an example, take the Schrödinger equation for a quantum scalar field:

$$i\hbar \frac{\partial \Psi(\phi, t)}{\partial t} = \int d^3x \frac{1}{2} \left[-\hbar^2 \frac{\delta^2}{\delta \phi^2} + (\nabla \phi)^2 + V(\phi) \right] \Psi(\phi, t) \quad (185)$$

Writing again $\Psi = R \exp(iS/\hbar)$, we obtain:

$$\frac{\partial S}{\partial t} + \int d^3x \frac{1}{2} \left[-\left(\frac{\delta S}{\delta \phi}\right)^2 + (\nabla \phi)^2 + V(\phi) \right] + Q = 0 \quad (186)$$

$$\frac{\partial R^2}{\partial t} + \int d^3x \frac{\delta}{\delta \phi} \left(R^2 \frac{\delta S}{\delta \phi} \right) = 0 \quad (187)$$

where $Q[\phi, t] = -\hbar^2 \frac{1}{2R} \int d^3x \frac{\delta^2 R}{\delta \phi^2}$ is the corresponding quantum potential.

A detailed analysis of the ontological interpretation of quantum field theory is given in Ref. [96] for the case of quantum electrodynamics.

The ontological interpretation of canonical quantum gravity is obtained in an analogous way. Substitution of $\Psi = R \exp(iS/\hbar)$ into the Wheeler-DeWitt equation (73) yields the two equations (for simplicity we stay in pure gravity):

$$\frac{1}{2} G_{ijkl} \frac{\delta S}{\delta h_{ij}} \frac{\delta S}{\delta h_{kl}} + h^{1/2} R^{(3)}(h_{ij}) + h^{1/2} Q(h_{ij}) = 0 \quad (188)$$

$$G_{ijkl} \frac{\delta}{\delta h_{ij}} \left(R^2 \frac{\delta S}{\delta h_{kl}} \right) = 0 \quad (189)$$

where the quantum potential is given by:

$$Q = -\hbar^2 \frac{1}{R} G_{ijkl} \frac{\delta^2 R}{\delta h_{ij} \delta h_{kl}} \quad (190)$$

As before, we set:

$$\Pi_{ij} = -h^{1/2} (K_{ij} - h_{ij} K) = \frac{\delta S}{\delta h^{ij}} \quad (191)$$

Hence, as K_{ij} is essentially the time derivative of h_{ij} , equation (191) gives the time evolution of h_{ij} , which will be different from the time evolution of classical general relativity due to the presence of the quantum potential in equation (188). As we see, there is no issue of time. The notion of spacetime is meaningful in this interpretation, exactly like the notion of trajectory is meaningful in particle quantum mechanics following this interpretation.

Some interesting works have been done in ontological quantum cosmology. In Ref. [97], it is shown, using a minisuperspace model, that the quantum potential cancels the cosmological constant for real wave functions (which is the case of the no-boundary one), leading to a null effective cosmological constant. In Ref. [98] it is shown, in a straightforward way, how to obtain semi-classical quantum cosmology, in the light of this interpretation.

7 Conclusion

In these lectures, we have tried to outline the main problems, goals, and achievements of quantum cosmology.

1) The first problem is to formulate a consistent theory of quantum gravity. In our opinion, perturbative methods are inadequate to quantize gravity. We have to focus our attention to a non-perturbative approach, which can be applied to some general theory which contains general relativity, like superstring theory [99], or simply to general relativity itself. We have adopted this last attitude, for simplicity, applied the Dirac quantization procedure, and arrived at the the fundamental canonical quantum gravity equation: the Wheeler-DeWitt equation (81), which is a complicate equation and very difficult to solve.

a) One attempt to solve this problem is to work with a different set of canonical variables: the Ashtekar variables [100, 101]. In these variables, the hamiltonian constraints of general relativity are greatly simplified, and also the corresponding Wheeler-DeWitt equation. They become similar to the Yang-Mills equations. Therefore, many technics developed for so many years in the framework of Yang-Mills theory, can be applied to canonical quantum gravity in Ashtekar variables. In particular, the introduction of loop-space variables, inspired in the Yang-Mills Wilson loops, has presented some interesting non-perturbative results [102, 103]. The program of canonical quantization of general relativity in the Ashtekar variables is the subject of intense research nowadays. It has many unresolved questions and problems of physical interpretation, but it surely deserves the attention and work of researchers interested in quantum gravity.

b) As we have mentioned in section 3, due to the Klein-Gordon nature of the Wheeler-DeWitt equation, it has been proposed that we should quantize the wave functions Ψ which are solutions of the Wheeler-DeWitt equation. It would be like a third quantization of gravity. This is a theory where universes, with all their internal characteristic features (total charge, coupling constants, etc) can be spontaneously created. This approach is connected with euclidean quantum gravity, where topology change is described in terms of instantons (solutions of the euclidean field equations, like in quantum field theory): disconnected universes (baby universes) are created by quantum fluctuations of topology [77]. In Refs. [104, 105] this approach is studied and the issue of the determination of the values of coupling constants is discussed.

2) In section 2 we have shown that the Copenhagen interpretation cannot be applied to quantum cosmology.

a) One proposal of interpretation is the one inspired in the many-worldss interpretation of quantum mechanics which says that a definite prediction can be made if the wave function has one peak. The problems of this interpretation are mainly two. First, it is not easy to find a peak in solutions of the Wheeler-DeWitt equation. One attempt to solve this problem is to study the corresponding Wigner functions of semi-classical wave functions but even them have not a single peak. Nevertheless, decoherence effects may yield a Wigner function with a single peak, but with a finite height. The natural question

is: what height must have a peak in order to be considered as a prediction? Furthermore, decoherence comes from tracing out irrelevant degrees of freedom. What are the irrelevant degrees of freedom in quantum cosmology?

b) A second proposal is the consistent histories interpretation. In this interpretation, the conditions for assigning probabilities to histories are established without mentioning observers or measurements.

A common feature of these two interpretations is the important role of decoherence in both of them. Decoherence is fundamental to obtain classical spacetime in a theory where there is no classical domain ‘a priori’.

c) The third interpretation presented in these lectures is fundamentally different from the first two. It is an ontological interpretation. In this approach, the notion of trajectories of quantum particles is meaningful. Analogously, the notion of spacetime is meaningful in quantum gravity and hence the notion of time. The classical limit is very easy to obtain; we have just to set the conditions for having the quantum potential equal to zero. The problem with this interpretation is the difficulty to accommodate the notion of spin, which cannot be described with classical images [106].

3) In order to obtain predictions from quantum cosmology, we need boundary conditions to the Wheeler-DeWitt equation which select only one of its solutions. In section 5, we have presented one of the proposed boundary conditions, the no-boundary one. The no-boundary wave function is analogous to a ground state wave function. We have shown, with a simple minisuperspace example, how a no-boundary wave function yields definite predictions. However, in general, the no-boundary boundary condition, does not select a unique wave function. It depends on the complex contour where we perform the path integral. Hence, it needs more specifications in order to yield a unique solution.

We have also described how quantum cosmology can be relevant in explaining structure formation. This is a domain where we can really test if quantum cosmology ideas may yield some physical testable consequences [107].¹⁶

Summarizing, a lot of work has been made and is still needed in domains like non-perturbative string theory, Ashtekar variables, baby universes, decoherence, consistent histories interpretation, Wigner functions, boundary conditions and the problem of structure formation. Not so much attention has been devoted, however, to the ontological interpretation which seems to be the best adapted to quantum cosmology, although with its problem to accommodate spin. We think some more research is needed in this area.

¹⁶Some possible effects of quantum gravity in quantum field theory have also been studied [108].

References

- [1] S. W. Hawking and G. F. R. Ellis; *The large scale structure of space-time* (Cambridge University Press, Cambridge, 1973).
- [2] C. W. Misner; Phys. Rev. **186**, 1319 (1969).
- [3] C. W. Misner; Phys. Rev. **186**, 1328 (1969).
- [4] B. S. DeWitt, Phys. Rev. **160**, 1113 (1967).
- [5] G. F. R. Ellis in the *Proceedings of the 6th Brazillian School on Cosmology and Gravitation*, ed. by M. Novello (World Scientific, New York, 1987).
- [6] S. W. Hawking; Phys. Rev. **D46**, 603 (1992).
- [7] K. S. Thorn in the *Proceedings of the 13th International Conference on General Relativity and Gravitation*, ed. by R. J. Gleiser, C. N. Kozameh and O. M. Moreschi (Institute of Physics Publishing, London, 1993).
- [8] R. M. Wald; *General Relativity* (The University of Chicago Press, Chicago, 1984).
- [9] M. Novello and J. M. Salim; Phys. Rev. **D20**, 377 (1979).
- [10] E. Elbaz, M. Novello, J. M. Salim and L. A. R. Oliveira; Int. J. of Mod. Phys. **D1**, 641 (1993).
- [11] G. F. R. Ellis, A. Sumeruk, D. Coule and C. Hellaby; Class. Quantum Grav. **9**, 1535 (1992).
- [12] J. V. Narlikar in the *Proceedings of the 6th Brazillian School on Cosmology and Gravitation*, ed. by M. Novello (World Scientific, New York, 1987).
- [13] G. Murphy; Phys. Rev. **D8**, 4231 (1973).
- [14] M. Mars; Phys. Rev. **D51**, 3989 (1995).
- [15] B. S. DeWitt in *Gravitation: An Introduction to Current Research*, ed. by L. Witten (Wiley, New York, 1962).
- [16] A. H. Guth; Phys. Rev. **D28**, 347 (1981).
- [17] E. W. Kolb and M. S. Turner; *The Early Universe* (Addison-Wesley Publishing Company, New York, 1990).
- [18] B. S. DeWitt in *The Many-Worlds Interpretation of Quantum Mechanics*, ed. by B. S. DeWitt and N. Graham (Princeton University Press, Princeton, 1973).
- [19] B. S. DeWitt; Physics Today **30** (september 1970).
- [20] C. Cohen-Tannoudji, B. Diu, F. Laloë, *Mécanique Quantique* (Hermann, Paris, 1977).

- [21] R. Omnès; *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, 1994).
- [22] H. Everett; *Rev. Mod. Phys.* **29**, 454 (1957).
- [23] *The Many-Worlds Interpretation of Quantum Mechanics*, ed. by B. S. DeWitt and N. Graham (Princeton University Press, Princeton, 1973).
- [24] F. Tipler; *Phys. Rep.* **137**, 231 (1986).
- [25] W. H. Zurek; *Physics Today* **36** (october 1991).
- [26] W. H. Zurek; *Phys. Rev.* **D24**, 1516 (1981).
- [27] W. H. Zurek; *Phys. Rev.* **D26**, 1862 (1982).
- [28] W. H. Zurek; *Phys. Rev.* **D40**, 1071 (1989).
- [29] A. O. Caldeira and A. J. Leggett; *Ann. Phys.* **149**, 374 (1983).
- [30] D. Bohm and B. J. Hiley; *The undivided universe: an ontological interpretation of quantum theory* (Routledge, London, 1993).
- [31] P. R. Holland; *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Interpretation of Quantum Mechanics* (Cambridge University Press, Cambridge, 1993).
- [32] M. H. Goroff and A. Sagnotti; *Nucl. Phys.* **B226**, 709 (1986).
- [33] K. Stelle; *Phys. Rev.* **D16**, 953 (1977).
- [34] D. Gross and V. Periwal; *Phys. Rev. Lett.* **60**, 2105 (1988).
- [35] E. Witten; *Nucl. Phys.* **B311**, 46 (1988).
- [36] C. Kiefer; *Phys. Rev.* **D45**, 2044 (1992).
- [37] A. H. Guth and S.-Y. Pi; *Phys. Rev.* **D32**, 1899 (1985).
- [38] J. N. Islam; *Found. of Phys.* **24**, 593 (1994).
- [39] C. W. Misner, K. S. Thorn and J. A. Wheeler; *Gravitation* (W. H. Freeman and Company, New York, 1973).
- [40] S. W. Hawking in *Euclidean Quantum Gravity*, ed. by G. W. Gibbons and S. W. Hawking (World Scientific, London, 1993).
- [41] T. Regge and C. Teitelboim; *Ann. Phys.* **88**, 286 (1974).
- [42] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Yeshiva University, New York, 1964).

- [43] M. Henneaux and C. Teitelboim; *Quantization of Gauge Systems* (Princeton University Press, Princeton, 1992).
- [44] K. Sundermeyer; “Constrained Dynamics” (Springer-Verlag, Berlin, 1982).
- [45] M. Gleiser, R. Holman and N. Pinto-Neto; Nucl. Phys. **B294**, 1164 (1987).
- [46] N. Pinto-Neto and A. F. Velasco; GRG **25**, 991 (1993).
- [47] K. Kuchar in *Quantum Gravity 2: A Second Oxford Symposium*, eds. C. J. Isham, R. Penrose and D. W. Sciama (Clarendon Press, Oxford, 1981).
- [48] A. Strominger; Phys. Rev. Lett **52**, 1733 (1984).
- [49] A. Hosoya and M. Morikawa; Phys. Rev. **D39**, 1123 (1989).
- [50] S. Giddings and A. Strominger; Nucl. Phys. **B321**, 481 (1989).
- [51] K. Ghoroku; Class. Quantum Grav. **8**, 447 (1991).
- [52] J. Barbour; Class. Quantum Grav. **11**, 2853 (1994).
- [53] J. Barbour; Class. Quantum Grav. **11**, 2875 (1994).
- [54] T. Banks; Nucl. Phys. **B249**, 332 (1985).
- [55] T. P. Singh and T. Padmanabhan; Ann. Phys. **196**, 296 (1989).
- [56] D. Giulini and C. Kiefer; Class. Quantum Grav. **12**, 403 (1995).
- [57] S. Tomonaga; Prog. Theor. Phys. **1/2**, 1 (1946).
- [58] R. Geroch; Noûs **18**, 617 (1984).
- [59] J. B. Hartle; Am. J. Phys **36**, 704 (1968).
- [60] W. Unruh; private communication. See also the paper by Neill Graham in Ref. [23].
- [61] M. V. Berry; Philos. Trans. R. Soc. London **287A**, 237 (1977).
- [62] J. J. Halliwell; Phys. Rev. **D36**, 3626 (1987).
- [63] M. Novello, V. M. C. Pereira and N. Pinto-Neto; Int. J. of Mod. Phys. **D4**, 673 (1995).
- [64] M. P. Ryan, lecture notes from the 6th *Brazillian School on Cosmology and Gravitation* (Rio de Janeiro, 1989).
- [65] K. V. Kuchar and M. P. Ryan; Phys. Rev. **D40**, 3982 (1989).
- [66] S. W. Hawking and D. N. Page; Nucl. Phys. **B264**, 185 (1986).
- [67] T. Christodoulakis and J. Zanelli; Class. Quantum Grav. **4**, 851 (1987).

- [68] I.N. Sneddon, *Elements of Partial Differential Equations* (McGraw-Hill Book Company, Inc., New York, 1957).
- [69] A. Anderson; Phys. Rev. **D42**, 585 (1990).
- [70] G. W. Gibbons and L. P. Grishchuk; Nucl. Phys. **B313**, 736 (1988).
- [71] A. D. Linde; Rep. Prog. Phys. **47**, 925 (1984).
- [72] W. M. Suen and K. Young; Phys. Rev. **D39**, 2201 (1989).
- [73] A. Vilenkin; Phys. Rev. **D33**, 350 (1986).
- [74] A. Vilenkin; Phys. Rev. **D37**, 888 (1988).
- [75] J. B. Hartle and S. W. Hawking; Phys. Rev. **D28**, 2960 (1983).
- [76] S. W. Hawking; Nucl. Phys. **B239**, 257 (1984).
- [77] *Euclidean Quantum Gravity*, ed. by G. W. Gibbons and S. W. Hawking (World Scientific, London, 1993).
- [78] J. J. Halliwell and J. Louko; Phys. Rev. **D39**, 2206 (1989); Phys. Rev. **D40**, 1868 (1989); Phys. Rev. **D42**, 3997 (1990).
- [79] J. J. Halliwell and J. B. Hartle; Phys. Rev. **D41**, 1815 (1990).
- [80] C. Teitelboim; Phys. Rev. **D25**, 3159 (1982).
- [81] W. A. Wright and I. G. Moss; Phys. Lett. **154B**, 115 (1985).
- [82] J.J. Halliwell in *Quantum Cosmology and Baby Universes*, ed. by S. Coleman, J.B. Hartle, T. Piran and S. Weinberg (World Scientific, Singapore, 1991).
- [83] J. J. Halliwell and S. W. Hawking; Phys. Rev. **D31**, 1777 (1985).
- [84] L. P. Grishchuk; Phys. Rev. Lett. **70**, 2371 (1993).
- [85] L. P. Grishchuk; Phys. Rev. **D50**, 7154 (1994).
- [86] S. Habib and R. Laflamme; Phys. Rev. **D42**, 4056 (1990).
- [87] R. B. Griffiths, Journal of St. Phys. **36**, 219 (1984).
- [88] M. Gell-Mann and J. B. Harle in *Complexity, Entropy and the Physics of Information*, ed. by W. Zurek (Addison Wesley, 1990).
- [89] J. B. Hartle in the *Proceedings of the 13th International Conference on General Relativity and Gravitation*, ed. by R. J. Gleiser, C. N. Kozameh and O. M. Moreschi (Institute of Physics Publishing, London, 1993).
- [90] A. M. Gleason; J. Math. Mechanics **6**, 895 (1953).

- [91] H. F. Dowker and J. J. Halliwell; *Phys. Rev.* **D46**, 1580 (1992).
- [92] C. Kiefer; *Class. Quantum Grav.* **8**, 379 (1991).
- [93] J. P. Paz and W. H. Zurek; *Phys. Rev.* **D48**, 2728 (1993).
- [94] C. Kiefer and H. D. Zeh; *Phys. Rev.* **D51**, 4145 (1995).
- [95] J. P. Paz and S. Sinha; *Phys. Rev.* **D44**, 1038 (1991).
- [96] P. N. Kaloyerou; *Phys. Rep.* **244**, 287 (1994).
- [97] E. J. Squires; *Phys. Lett.* **A162**, 35 (1992).
- [98] J. Kowalski-Glikman and J. C. Vink; *Class. Quantum Grav.* **7**, 901 (1990).
- [99] M. Green, J. Schwarz and E. Witten; *Superstring Theory* (Cambridge, Cambridge University Press, 1987).
- [100] A. Ashtekar; *Phys. Rev.* **D36**, 1587 (1987).
- [101] J. Samuel; *Pramana J. Phys.* **28**, L429 (1987).
- [102] C. Rovelli and L. Smolin; *Nucl. Phys.* **B331**, 80 (1990).
- [103] A. Ashtekar; in *Non-Perturbative Canonical Gravity*, IUCAA Lecture Notes (Syracuse, New York, 1990).
- [104] S. Coleman; *Nucl. Phys.* **B310**, 643 (1988).
- [105] W. Fischler, I. Klebanov, J. Polchinski and L. Susskind; *Nucl. Phys.* **B327**, 157 (1989).
- [106] H. C. Ohanian; *Am. J. Phys.* **54**, 500 (1986).
- [107] L. P. Grishchuk in the *Proceedings of the 6th Brazillian School on Cosmology and Gravitation*, ed. by M. Novello (Editions Frontieres, Gif-sur-Yvette, France 1994).
- [108] E. Elizalde, S. D. Odintsov and A. Romeo; *Phys. Rev.* **D51**, 4250 (1995).