The strange equation of quantum gravity

Carlo Rovelli

Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France.

E-mail: rovelli@cpt.univ-mrs.fr

Received 28 September 2014, revised 22 October 2014
Accepted for publication 24 October 2014
Published 1 June 2015

Abstract
Disavowed by one of its fathers, ill-defined, never empirically tested, the Wheeler–DeWitt equation has nevertheless had a powerful influence on fundamental physics. A well-deserved one.

Keywords: quantum gravity, Wheeler–DeWitt equation, canonical general relativity

1. Introduction

One day in 1965, John Wheeler had a two-hour stopover between flights at the Raleigh–Durham airport in North Carolina. He called Bryce DeWitt, then at the University of North Carolina in Chapel Hill, proposing to meet at the airport during the wait. Bryce showed up with the Hamilton–Jacobi equation of general relativity, published by Asher Peres a little earlier [1]:

\[
\frac{\delta S[q]}{\delta q_{ab}} - \frac{\delta S[q]}{\delta q_{cd}} \frac{\delta S[q]}{\delta q_{ac}} = 0.
\]

The \( q_{ab} \), with \( a, b = 1, 2, 3 \), are the components of a 3D metric, \( R[q] \) its Ricci scalar, and \( S[q] \) a Hamilton–Jacobi functional. Bryce mumbled the idea of repeating what Schrödinger did for the hydrogen atom: getting a wave equation by replacing the square of derivatives with \((i)\) times a second derivative—a way of undoing the optical approximation. This gives [2, 3]

\[
\left( q_{ab} q_{cd} - \frac{1}{2} q_{ac} q_{bd} \right) \frac{\delta}{\delta q_{ac}} \frac{\delta}{\delta q_{bd}} + \det q \ R[q] = 0.
\]

for a functional \( \Psi[q] \) to be interpreted as the wavefunction of the gravitational field \( q \). Wheeler got tremendously excited (he was often enthusiastic) and declared on the spot that the equation for quantum gravity had been found. This is the birth of equation (2) ([4], p 58).

For a long time Wheeler called it the ‘Einstein–Schrödinger equation’, while DeWitt called it
the ‘Wheeler equation’, or ‘that damn equation’, not hiding his unhappiness with it. For the rest of the world it is the Wheeler–DeWitt, or ‘WdW’, equation.

It is a strange equation, full of nasty features. First, it is ill-defined. The functional derivatives are distributions, which cannot be squared without yielding divergences. Concrete calculations, indeed, tend to give meaningless results: strictly speaking, there is no equation. Second, the equation breaks the manifest relativistic covariance of space and time: quite bad for the quantum theory of general relativity. Third, there is no time variable in the equation, a puzzling feature for an equation with the ambition to be the dynamical equation of quantum spacetime. This has raised endless confusion, and prompted a long lasting debate on the nature of time.

Furthermore, a Schrödinger-like equation is not sufficient for defining a quantum theory: a scalar product is needed to compute expectation values. A scalar product suitable for the WdW equation is far from obvious, and rivers of ink have been wasted on discussing this issue as well. Last, but obviously not least, half a century after the publication of the equation, none of the numerous tentative predictions derived from it have yet been verified or supported by observation.

Still, in spite of all this, the WdW equation is a milestone in the development of general relativity. It has inspired a good part of the research in quantum gravity for decades, and has opened new perspectives for fundamental physics. Much of the original conceptual confusion raised by the equation has been clarified. Consensus on a theory of quantum gravity is lacking, but tentative theoretical constructions, such as strings and loops, exist today. Well-defined versions of the equation have been found, and they are utilized to produce predictions, with a hope of testing them against observation. The equation is recognized as a basic tool for thinking about the quantum properties of spacetime by a large community of physicists in areas ranging from early cosmology to black hole physics. The equation has proven an inexhaustible source of inspiration on the path towards understanding the quantum properties of space and time.

The main reason for this is that the WdW equation has been the icebreaker in working towards the construction of a quantum theory which does not presuppose a single spacetime. The physics community underwent a major reshaping of its way of thinking when it began to utilize general relativity: from a conception of physics as the theory of what happens in space and time to a new conception, where the theory also describes what happens to space and time. This step has been difficult in the classical framework, where questions like that of whether gravitational waves were physical or gauge, or concerning the nature of the Schwarzschild singularity, have confused the community for decades. But it has taken longer, and in fact is still confusing the community, for the quantum regime. Quantum mechanics and general relativity, taken together, imply the possibility of quantum superposition of different spacetimes. The WdW equation, which is based on a wavefunction $\Psi[q]$ over geometries, has offered the first conceptual scheme for dealing with this physical possibility. For this reason it is a milestone.

Today’s tentative quantum gravity theories take for granted the absence of a single predetermined smooth spacetime at all scales. The breakthrough opening the path to this thinking happened at the Raleigh–Durham airport 1965 meeting.

2. From Einstein–Hamilton–Jacobi to Einstein–Schrödinger

In the following I make no attempt at a full review of the vast literature wherein the WdW equation has been discussed, has been utilized, and has left its mark. Instead, I focus on the
main issues that the equation has raised, on what we have learned from it, and on the possibility that it has opened for describing quantum spacetime.

I start by illustrating the path leading to the WdW equation via Peres’s Hamilton–Jacobi formulation of general relativity [1], in a bit more detail than above. This is relevant not only for history, but also, and especially, because it clarifies the meaning of the WdW equation, shedding light on the confusion that it raised. Some difficulties of the WdW equation are more apparent than real. They have nothing to do with quantum mechanics; they stem from the application of the Hamilton–Jacobi language to a generally covariant context.

Einstein wrote general relativity in terms of the Lorentzian metric $g_{\mu\nu}(\vec{x}, t)$, where the $\mu, \nu = 0, \ldots, 3$ are spacetime indices and the $(\vec{x}, t)$ are coordinates in spacetime. Peres’s starting point was the Arnowit–Deser–Misner change of variables [5]

\[ q_{ab} = g_{ab}, \quad N^a = g^{ab}, \quad N = \sqrt{-g^{00}}, \]

where $a, b = 1, 2, 3$. $N$ and $N^a = q^{ab}N_b$ are the ‘Lapse’ and ‘Shift’ functions (respectively a scalar and a vector), and $q^{ab}$ is the inverse of $q_{ab}$. The interest of these variables is that the Lagrangian does not depend on the time derivatives of Lapse and Shift. Therefore the only true dynamical variable is the 3-metric $q_{ab}(\vec{x})$, the Riemannian metric of a $t = \text{constant}$ surface. The canonical Hamiltonian vanishes, as is always the case for systems invariant under reparameterization of the Lagrangian evolution parameter, and the dynamics is coded by the two ADM constraints

\[ \left( q_{ab}q_{cd} - \frac{1}{2}q_{ac}q_{bd} \right) p^{ac}p^{bd} - \det q R[q] = 0 \]

and

\[ D_a p^{ab} = 0, \]

on the 3-metric $q$ and its conjugate momentum $p$. ($D_a$ is the covariant derivative of the 3-metric.)

The dynamics can be written in Hamilton–Jacobi form by introducing the Hamilton–Jacobi function $S[q]$, which is a functional of $q_{ab}(\vec{x})$, and demanding that this satisfies the two equations above with $p^{ab}$ replaced by the functional derivative $\delta S[q]/\delta q_{ab}$. The first equation gives (1), while the second can be rather easily shown [6] to be equivalent to the requirement that

\[ S[q] = \tilde{S}[\tilde{q}] \]

for any two 3D metrics $q$ and $\tilde{q}$ related by a 3D change of coordinates. Solving (1) and (6) amounts to solving the Einstein equations. To see this, say we have found a family of solutions $S[q, q^0]$ parameterized by a 3-metric $q^0_{ab}(\vec{x})$. Then we can define the momenta

\[ p_0^{ab}[q, q^0](\vec{x}) = \frac{\delta S[q, q^0]}{\delta q^0_{ab}(\vec{x})} \]

which are thus nonlocal functions of $q$ and $q^0$. For any choice of $q^0$ and $p_0$ satisfying (4) and (5), there exists a spacetime which is a solution of Einstein’s equations and for any such solution, any field $q$ satisfying

\[ p_0^{ab}[q, q^0](\vec{x}) = p_{0}^{ab}(\vec{x}) \]
is the metric of a spacelike 3D surface imbedded in this spacetime. Since the $q^0$ and $p_0$ are related by equations (4) and (5), the last equation does not determine $q$ uniquely: the different solutions correspond to different spacelike surfaces in spacetime, and different coordinatization of the same; in fact, all of them. Therefore the solution to the Einstein–Hamilton–Jacobi system (1)–(6) provides in principle the full solution of Einstein’s equations. These results follow from a simple generalization of standard Hamilton–Jacobi theory.

This was the starting point of Bryce DeWitt. Now recall that in his milestone 1926 article [7], Schrödinger introduced (what is called today) the Schrödinger equation for the hydrogen atom by taking the Hamilton–Jacobi equation of an electron in a Coulomb potential:

$$\frac{\partial S(\vec{x}, t)}{\partial t} + \frac{1}{2m} \frac{\partial S(\vec{x}, t)}{\partial \vec{x}} \cdot \frac{\partial S(\vec{x}, t)}{\partial \vec{x}} + \frac{e^2}{|\vec{x}|} = 0$$

and replacing derivatives with $(-i\hbar)$ times derivative operators:

$$\left[ -i\hbar \frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial \vec{x}^2} + \frac{e^2}{|\vec{x}|} \right] \psi(\vec{x}, t) = 0.$$

In an article presented shortly after [8], Schrödinger offers a rationale for this procedure: the eikonal approximation of a wave equation, which defines the geometrical optic approximation where wave packets follow definite trajectories, can be obtained by the opposite procedure. If we interpret classical mechanics as the eikonal approximation to a wave mechanics, we can guess the wave equation by this procedure. Given the immense success of Schrödinger’s leap, trying the same strategy for gravity is obviously tempting. This is what led DeWitt and Wheeler to equation (2). Next to it, the second Hamilton–Jacobi equation, following from (5), remains unaltered:

$$D_a \frac{\delta}{\delta q^a} \Psi[q] = 0,$$

and, as before, is equivalent to the requirement that $\Psi[q] = \Psi[\tilde{q}]$ if $q$ and $\tilde{q}$ are related by a 3D coordinate transformation. That is, the wavefunction is only a function of the ‘3-geometry’, namely the equivalence class of metrics under a diffeomorphism, and not of the specific coordinate-dependent form of the $g_{ab}(\vec{x})$ tensor.

The Schrödinger equation (10) gives, in a sense, the full dynamics of the electron in the hydrogen atom. Similarly, one expects the WdW equation (2), properly understood and properly defined, to give the full dynamics of quantum gravity.

3. Physics without background time

The immediately puzzling aspect of the WdW equation, and the one that has raised the largest confusion, is the absence of a time variable in the equation. This has often been wrongly attributed to some mysterious quantum disappearance of time. But things are simpler: the disappearance of the time variable is already a feature of the classical Hamilton–Jacobi formulation of general relativity. It has nothing to do with quantum mechanics. It is only a consequence of the peculiar manner in which evolution is described in general relativity.

In Newtonian and special relativistic physics, the time variable represents the reading of a clock. It is therefore a quantity with which we associate a well-determined procedure of measurement. Not so in general relativity, where the reading of a clock is not given by the time variable $t$, but is instead expressed by a line integral depending on the gravitational field,
computed along the clock’s worldline $\gamma$:

$$T_\gamma = \int \sqrt{g_{\mu\nu} \, dx^\mu \, dx^\nu}. \quad (12)$$

The coordinate $t$ in the argument of $g_{\mu\nu}(\vec{x}, \, t)$, which is the evolution parameter of the Lagrangian and Hamiltonian formalisms, has no direct physical meaning and can be changed freely. Such change in the manner in which evolution is described is not a minor step. Einstein wrote that the biggest difficulty he had to overcome in order to find general relativity was to understand ‘the meaning of the coordinates’\(^1\). The physical predictions of general relativity, which can be directly tested against experience, are not given by the evolution of physical quantities in the coordinate $t$, but, rather, by the relative evolution of physical quantities (among which are the proper times of equation (12)) with respect to one another. This is why using the time variable $t$ is not required for making sense of general relativity.

To clarify this crucial point, consider two clocks on the surface of the Earth; imagine that one of them is thrown upward, and then falls back down near the first. The readings of the two clocks, say $T_1$ and $T_2$, initially the same, will then differ. Given the appropriate initial data, general relativity allows us to compute the value of $T_1$ when the second clock reads $T_2$, or vice versa. Does this describe the evolution of $T_1$ in the ‘time’ $T_2$, or, instead, the evolution of $T_2$ in the ‘time’ $T_1$?

The question is clearly pointless: general relativity describes the relative evolution of the two variables $T_1$ and $T_2$, both given by (12) but computed along different worldlines. The two ‘times’ $T_1$ and $T_2$ are on the same footing. The example shows that general relativity describes the relative evolution of variable quantities with respect to one another, and not the absolute evolution of variables in time.

Mathematically this is realized by parameterizing the motions. Instead of using $T_1(T_2)$ or $T_2(T_1)$ to describe evolution, the theory uses the parametric form $T_1(t)$, $T_2(t)$, where $t$ is an arbitrary parameter, which can be chosen freely. The coordinates $(t, \vec{x})$ in the argument of the gravitational field $g_{\mu\nu}(\vec{x}, \, t)$ are arbitrary parameters of this sort.

This manner of describing evolution is more general than giving the evolution in a preferred time parameter. The formal structure of dynamics can be generalized to this wider context. This was recognized early on by Dirac [10], and the corresponding generalized formulation of mechanics has been discussed by many authors in several variants (see for instance [11, 12]), often under the deceptively restrictive denomination of ‘constrained system dynamics’. ‘Constrained system dynamics’ is not the dynamics of special systems that have constraints: it is a generalization of dynamics which avoids the need of picking one of the variables and treating it as the special, independent, evolution parameter. In the corresponding generalization of Hamilton–Jacobi theory, parameter time does not appear at all, because the Hamilton–Jacobi theory gives the relation between observable variables directly.

Thus, the reason that the coordinate time variable $t$ does not show up in the WdW equation is not mysterious after all. It is the same reason for which coordinates do not show up in the physical predictions of classical general relativity: they have no physical meaning, and the theory can well do without them.

In particular, the absence of $t$ in the WdW equation does not imply that the theory describes a frozen world, as unfortunately often suggested. One can pick a function of the gravitational field, or, more realistically, couple a simple system to the gravitational field, and

---

\(^1\) ‘Why were a further seven years required for setting up the general theory of relativity? The principal reason is that one does not free oneself so easily from the conception that an immediate physical significance must be attributed to the coordinates.’ Albert Einstein, in [9].
use it as a physical clock, to coordinatize evolution in a physically relevant manner. A common strategy in quantum cosmology, for instance, is to include a scalar field $\phi(\vec{x}, t)$ in the system studied, take the approximation where $\phi(\vec{x}, t)$ is constant in space, $\phi(\vec{x}, t) = \phi(t)$, and give it a simple dynamics, such as a linear growth in proper time. Then the value of $\phi$ can be taken as a ‘clock’—it coordinatizes trajectories of the system—and the WdW wavefunction $\Psi[q, \phi]$ can be interpreted as describing the evolution of $\Psi[q]$ in the physical clock variable $\phi$, tuned to proper time.

The full structure of quantum mechanics in the absence of a preferred time variable has been studied by several authors. Reference [12] is the favorite version of this author. It can be synthesized as follows. The variables that can interact with an external system (‘partial observables’) are represented by self-adjoint operators $A_n$ in an auxiliary Hilbert space $\mathcal{K}$ where the WdW operator $C$ is defined. The WdW equation

$$C\Psi = 0 \quad (13)$$

defines a linear subspace $\mathcal{H}$, and we call $P: \mathcal{H} \to \mathcal{H}$ the orthogonal projector. If zero lies in the continuous spectrum of $C$, then $\mathcal{H}$ is not a proper subspace of $\mathcal{K}$. It is a generalized subspace, namely a linear subspace of a suitable completion of $\mathcal{K}$ in a weak norm. In this case, $\mathcal{H}$ still inherits a scalar product from $\mathcal{K}$: this can be defined using various techniques, such as spectral decomposition of $\mathcal{K}$, or group averaging. $P: \mathcal{H} \to \mathcal{H}$ is still defined, and it is not a projector, as $P^2$ diverges, but we still have $CP = PC = 0$. If $A_n$ is a complete set of commuting operators (in the sense of Dirac) and $\langle n | A | a \rangle$ a basis diagonalizing them, then

$$W(a'; a) = \langle a' | P | a \rangle \quad (14)$$

is the amplitude for measuring $\{a_n'\}$ after $\{a_n\}$ has been measured, from which transition probabilities can be defined by properly normalizing.

This formalism is a generalization of standard quantum mechanics. It reduces to the usual case if

$$C = i\hbar \frac{\partial}{\partial t} + H \quad (15)$$

where $H$ is a standard Hamiltonian. In this case, if $\{q, t\}$ is a complete set of quantum numbers, then

$$W(q', t'; q, t) = \langle q', t' | P | q, t \rangle \quad (16)$$

turns out to be the standard propagator

$$W(q', t'; q, t) = \langle q' | e^{-\frac{i}{\hbar}H} | q \rangle \quad (17)$$

which has the entire information about the quantum dynamics of the system. Thus, in general the WdW equation is just a generalization of the Schrödinger equation, to the case where a preferred time variable is not singled out. Equation (2) is the concrete form that (13) takes when the dynamical system is the gravitational field.

The first merit of the equation written by Wheeler and DeWitt in 1965 is therefore that it has opened the way to the generalization of quantum theory needed for understanding the quantum properties of our general covariant world.

The world in which we live is, as far as we understand, well described by a general covariant theory. In principle, any quantity that we can observe and measure around us can be represented by an operator $A_n$ on $\mathcal{K}$, and all dynamical relations predicted by physics can be expressed in terms of the transition amplitudes (14). In practice, setting up a concrete realization of this framework is complicated. To see where we are today along the path of making
sense of the quantum theory formally defined by the WdW equation, and to describe the legacy of the equation, I now turn to some specific current approaches to quantum gravity.

4. WdW in strings

String theory developed starting from conventional quantum field theory and got in touch with general relativity only later. Even in dealing with gravity, string theory was for some time confined to perturbations around Minkowski space, where the characteristic features of general relativity are not prominent. For this reason, the specific issues raised by the fact that spacetime is dynamical have not played a major role in the first phases of development. But major issues cannot remain hidden long, and at some point string theory has begun to face dynamical aspects of spacetime. In recent years, the realization that in the world there is no preferred spacetime has impacted string theory substantially, and the string community has even gone to the opposite extreme, largely embracing more or less precise holographic ideas, where bulk spacetime is no longer a primary ingredient. In this context the WdW equation has reappeared in various ways in the context of the theory.

For instance, in the AdS/CFT setting, a constant radial coordinate surface can be seen as playing the role of the ADM constant time surface, the quantization of the corresponding constraint of the bulk gravity theory gives a WdW equation and its Hamilton–Jacobi limit turns out to admit an interpretation as a renormalization group equation for the boundary CFT [13, 14].

In fact, the full AdS/CFT correspondence can be seen as a realization of a WdW framework: the correspondence critically relies on the ADM Hamiltonian being a pure boundary term, since its ‘bulk’ part is pure gauge. See [15] and, for a direct attempt to derive the AdS/CFT correspondence from the WdW equation, see [16]. See also the discussion on dS/CFT, where the ‘wavefunction of the universe’ seen as a functional of 3-metrics, just as in canonical general relativity, plays a major role [17].

There is a more direct analog to the WdW equation in the foundation of string theory. In its ‘first quantization’, string theory can be viewed as a two-dimensional field theory on the world sheet. The action can be taken to be the Polyakov action, which is generally covariant on the world sheet, and therefore the dynamics is entirely determined by the constraints as in the case of general relativity. In fact, the two general relativity constraints (4) and (5) have a direct analog in string theory as the Virasoro constraints

\[ \Pi^2 - |VX|^2 = 0 \]  
and

\[ VX \cdot \Pi = 0, \]  

where \( X \) stands for the coordinates of the string and \( \Pi \) its momentum. As in general relativity, the second equation implements the invariance under a one-dimensional spatial change of coordinates on the world sheet, while the first is a ‘Hamiltonian constraint’ which codes its dynamics. The left-moving and right-moving null combinations (associated with \( x \pm t \)) are the left-moving and right-moving Virasoro algebras. However, the standard quantization of the string is obtained differently from how the WdW is derived. The two equations combined in right and left null combinations and Fourier transformed give the Virasoro operators \( L_n \), of which only the \( n \geq 1 \) components are imposed as operator equations on the quantum states. The expectation value of the \( L_n \) for negative \( n \) vanishes on physical states, so the full constraint is still recovered in the classical limit (in a similar way to the Gupta–Bleuler form).
In addition, states are required to be eigenstates of $L_0$, with eigenvalues given by the central charge of the corresponding conformal field theory.

The reason for these choices is that the theory is required to make sense in the target space and respect its Poincaré invariance. But imposing all the constraints $L_n = 0$ would be inconsistent, because of the nontriviality of the $L_n$ algebra: a warning about naive treatment of the Hamiltonian constraints that must also be kept in mind in the case of general relativity.

5. WdW in loops

The line of research where the WdW equation has had the largest impact is that of loop quantum gravity (LQG). LQG was in fact born from a set of solutions of the WdW equation, and its structure is still based on these solutions. General relativity can be rewritten in terms of variables different from those of the metric used by Einstein. These, developed by Abhay Ashtekar in the late 1980s, are the variables of an $SU(2)$ gauge theory, namely (in the Hamiltonian framework) an $SU(2)$ connection $A_a$ and its ‘electric field’ conjugate momentum $E^a$ [18]. Rewritten in terms of these variables, the Hamiltonian constraint of general relativity reads

$$C = F_{ab}E^aE^b = 0$$

and the WdW equation takes the simpler form

$$F_{ab} \frac{\delta}{\delta A_a} \frac{\delta}{\delta A_b} \Psi [A] = 0.$$  \hspace{1cm} (21)

Remarkably, we know a large number of solutions of this equation. These were first discovered using a lattice discretization by Ted Jacobson and Lee Smolin [19], and can be constructed, in the continuum, as follows. Choose a loop $\gamma = S_i \to R^3$, namely a closed line in space, and consider the trace of the holonomy of $A$ along this loop, namely the quantity

$$\int_\gamma \Psi = \gamma \int_\gamma P e^{\int A},$$

where $P$ indicates the standard path-ordered exponentiation. It turns out that $\Psi [A]$ is a solution of the Ashtekar–WdW equation (21) if the loop has no self-intersection (if $\gamma$ is injective) [20, 21]. A simplified way of deriving this result is to observe that the functional derivative of a holonomy vanishes at all space points $\vec{x}$ outside the loop, and is otherwise proportional to the tangent $\gamma^a = d\gamma^a(s)/ds$ to the loop:

$$\frac{\delta \Psi [A]}{\delta A_a(x)} = \int_\gamma ds \gamma^a(s) \delta^3(\gamma (s), \vec{x}) \left[ \tau_i P e^{\int A} \right].$$

where the path-ordered integral starts at the loop point $\vec{x}$, and $A_a = A_a^i \tau_i$, with $\tau_i, i = 1, 2, 3$, a basis in the $su(2)$ algebra. The left-hand side of (21) is therefore proportional to $F_{ab} \gamma^a \gamma^b$, which vanishes because of the antisymmetry of $F_{ab}$ if the loop has no self-intersection. If it has intersections, at the intersection point there are two different tangents and mixed terms do not cancel. Thus, loop states $\Psi$ without intersections are exact solutions of the WdW equation.

Nonintersecting loop states alone do not describe a realistic quantum space, because they are eigenstates of the volume $V = \det q_{ab}$ with vanishing eigenvalue. (Indeed, $V^2 \sim \epsilon_{abc}E^aE^bE^c \sim \epsilon_{abc}\gamma^{a\beta}\gamma^{b\gamma}\gamma^{c\delta} = 0$.) Therefore intersections play a role in the theory [22]. But acting on a loop state with intersections, the Ashtekar–WdW operator acts nontrivially only at the intersection point. This is the basic fact underpinning LQG.
The loop representation of quantum general relativity can formally be obtained by moving from the connection basis $\Psi[A] = \langle A|\Psi \rangle$ to a basis formed from loop states with intersections. An orthonormal basis of (linear combinations of) loop states with intersections is given by the spin network states [23], providing today the standard basis on which the theory is defined. In this basis, the WdW operator acts only at intersections, which are called the ‘nodes’ of the network. A cumbersome but well-defined definition of a WdW operator in this representation has been given by Thomas Thiemann, and is constructed and studied in detail in his book [24].

Simplified versions of this operator are used in loop quantum cosmology, the application of loop gravity results to quantum cosmology [25]. Loop quantum cosmology is producing tentative preliminary predictions about possible early universe quantum gravity effects on the CMB (see for example [26]). If these are verified, the WdW equation may turn out to be the key ingredient for the first successful predictions of quantum gravity.

6. WdW and path integrals

Since its earliest days [27], the search for a quantum theory of quantum gravity has alternatively looked for inspiration in the canonical WdW framework and in the covariant framework provided by a ‘path integral over geometries’:

$$Z = \int \mathcal{D}[g] e^{i\int \sqrt{g} R[g]}.$$  \hspace{1cm} (24)

It is hard to give this integral a mathematical sense, or to use it for computing transition amplitudes within some approximation scheme, but the formal expression (24) has provided another source of intuitive guidance for constructing the theory. Formally, the path integral is related to the WdW equation, in same manner in which the Feynman path integral that defines the propagator of a nonrelativistic particle is a solution of the Schrödinger equation. This relation has taken a particularly intriguing form in the context of the Euclidean quantum gravity program, developed by Hawking and his collaborators [28], where the wavefunctional

$$\Psi[q] = \int_{\partial q} \mathcal{D}[g] e^{-i\int \sqrt{g} R[g]}$$  \hspace{1cm} (25)

can be shown to be a formal solution of the WdW equation. Here the integration is over Euclidean 4-metrics, inducing the 3-metric $q$ on a 3D boundary. The construction lies at the heart of beautiful ideas such as the Hartle–Hawking ‘no boundary’ definition of a wavefunction for cosmology [29, 30].

The alternative use of canonical and covariant methods is not peculiar to quantum gravity. In fact, it is common in theoretical physics. The two approaches have complementary strengths: the Hamiltonian theory captures aspects that are easily overlooked in the covariant language, especially in the quantum context, while the Lagrangian framework allows symmetries to remain manifest, is physically more transparent, and leads more easily to calculation techniques.

In the loop context, the difficulties of dealing with the WdW equation have pushed a part of the community to adopt alternative covariant methods for computing the transition amplitudes (14). These make use of the so-called ‘spinfoam’ techniques [31, 32], a sum-over-paths technique for computing amplitudes connecting spin network states which can be seen as a well-defined version of equation (24).

After all, therefore, the initial unhappiness of Bryce DeWitt with an equation engulfed into the complexities of the Hamiltonian formalism and having the bad manners to break the manifest covariance of space and time were not unmotivated. The WdW equation does not
necessarily provide the best way of actually defining the quantum dynamics and computing transition amplitudes in quantum gravity.

But it remains the equation that has opened up the world of background independent quantum gravity, a unique source of inspiration, and a powerful conceptual tool that has forced us to understand how to actually make sense of a theory where space and time are quantum dynamical entities.

References

[1] Peres A 1962 On Cauchy’s problem in general relativity Nuovo Cimento 26 53
[2] DeWitt B S 1967 Quantum theory of gravity. 1. The canonical theory Phys. Rev. 160 1113–48
[3] Wheeler J A 1968 Superspace and the nature of quantum geometrodynamics Batailles Rencontres ed B S DeWitt and J A Wheeler (New York: Benjamin)
[4] DeWitt-Morette C 2011 The Pursuit of Quantum Gravity (Berlin: Springer-Verlag)
[5] Arnowitt R L, Deser S and Misner C W 1962 The dynamics of general relativity Gravitation: Introduction to Current Research p 227
[6] Higgs P W 1958 Integration of secondary constraints in quantized general relativity Phys. Rev. Lett. 1 373–4
[7] Schrödinger E 1926 Quantisierung als eigenwertproblem (Erste Mitteilung) Ann. Phys., Lpz. 79 361–76
[8] Schrödinger E 1926 Quantisierung als eigenwertproblem (Zweite Mitteilung) Ann. Phys. Lpz. 79 489–527
[9] Einstein A 1949 Autobiographical notes Albert Einstein: Philosopher-Scientist ed P Schilpp (London: Open Court: La Salle/Cambridge University Press) pp 2–94
[10] Dirac P A M 1958 Generalized Hamiltonian dynamics Proc. R. Soc. Lond. 246 326–32
[11] Hanson A, Regge T and Teitelboim C 1976 Constrained Hamiltonian Systems (Roma: Accademia Nazionale dei Lincei)
[12] Rovelli C 2004 Quantum Gravity. (Cambridge: Cambridge University Press)
[13] Lifschytz G and Periwal V 2000 Schwinger-Dyson = Wheeler-DeWitt: gauge theory observables as bulk operators J. High Energy Phys. JHEP00(2000)026
[14] Papadimitriou I and Skenderis K AdS/CFT correspondence and geometry Proceedings of the Strasbourg meeting on AdS/CFT arXiv:hep-th/0404176
[15] Marolf D 2009 Unitarity and holography in gravitational physics Phys. Rev. D 79 044010
[16] Freidel L Reconstructing AdS/CFT arXiv:0804.0632
[17] Maldaclena J 2003 Non-gaussian features of primordial fluctuations in single field inflationary models J. High Energy Phys. JHEP03(2003)013
[18] Ashtekar A 1986 New variables for classical and quantum gravity Phys. Rev. Lett. 57 2244–7
[19] Jacobson T and Smolin L 1988 Nonperturbative quantum geometries Nucl. Phys. B 299 295
[20] Rovelli C and Smolin L 1988 Knot theory and quantum gravity Phys. Rev. Lett. 61 1155
[21] Rovelli C and Smolin L 1990 Loop space representation of quantum general relativity Nucl. Phys. B331 80
[22] Brügmann B and Pullin J 1991 Intersecting N loop solutions of the Hamiltonian constraint of quantum gravity Nucl. Phys. B 363 221–44
[23] Rovelli C and Smolin L 1995 Spin networks and quantum gravity Phys. Rev. D 52 5743–59
[24] Thiemann T 2007 Modern Canonical Quantum General Relativity (Cambridge: Cambridge University Press)
[25] Ashtekar A 2013 Quantum gravity and quantum cosmology Lect. Notes Phys. 893 31–56
[26] Agullo I, Ashtekar A and Nelson W 2012 A quantum gravity extension of the inflationary scenario Phys. Rev. Lett. 109 251301
[27] Misner C W 1957 Feynman quantization of general relativity Rev. Mod. Phys. 29 497
[28] Hawking S W 1980 The path integral approach to quantum gravity General Relativity. An Einstein Centenary Survey ed S W Hawking and W Israel (Cambridge: Cambridge University Press) pp 746–89
[29] Hartle J B and Hawking S W 1983 Wave function of the universe Phys. Rev. D 28 2960–75
[30] Halliwell J J and Hawking S W 1985 The origin of structure in the universe Phys. Rev. D 31 1777
[31] Rovelli C and Vidotto F 2014 Introduction to Covariant Loop Quantum Gravity (Cambridge: Cambridge University Press)
[32] Perez A 2013 The spin-foam approach to quantum gravity Liv. Rev. Relativ. 16 3