IC/90/312

INTERNATIONAL CENTRE FOR
THEORETICAL PHYSICS

QUANTUM GRAVITATIONAL CORRECTIONS TO THE FUNCTIONAL SCHRÖDINGER EQUATION

Claus Kiefer

and

Tejinder P. Singh

INTERNATIONAL ATOMIC ENERGY AGENCY

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

1990 MIRAMARE - TRIESTE

IC/90/312

International Atomic Energy Agency
and
United Nations Educational Scientific and Cultural Organization
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

QUANTUM GRAVITATIONAL CORRECTIONS TO THE FUNCTIONAL SCHRÖDINGER EQUATION *

Claus Kiefer
Institut für Theoretische Physik der Universität Zürich,
CH-8001 Zürich, Switzerland

and

Tejinder P. Singh **
International Centre for Theoretical Physics, Trieste, Italy.

ABSTRACT

We derive corrections to the Schrödinger equation which arise from the quantization of the gravitational field. This is achieved through an expansion of the full functional Wheeler-DeWitt equation with respect to powers of the Planck mass. We demonstrate that the correction terms are independent of the factor ordering which is chosen for the gravitational kinetic term. Although the corrections are numerically extremely tiny, we show how they lead, at least in principle, to shift in the spectral lines of hydrogen type atoms. We discuss the significance of these corrections for quantum field theory near the Planck scale.

MIRAMARE – TRIESTE
October 1990

* Submitted for publication.
** Permanent address: Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400005, India.

1 Introduction

One of the many approaches towards a quantum theory of gravity is the application of canonical quantisation procedures to general relativity. In the Dirac approach, the four constraints which are present in the classical theory are implemented in the quantum theory through four conditions on wave functionals. The traditional, geometrodynamical, choice of variables leads to the concept of superspace – the configuration space of all (spacelike) three-geometries and matter fields. While three of the quantum constraints only express the invariance of the wave functional with respect to diffeomorphisms on three-space, the fourth one – the Wheeler-DeWitt equation – is responsible for the dynamics. It reads explicitly

$$\Big(-\frac{16\pi G\hbar^2}{c^2}G_{ijkl}\frac{\delta^2}{\delta h_{ij}\delta h_{kl}}-\frac{c^4}{16\pi G}\sqrt{h}\,{}^{(3)}R + h_m(h_{ij}(\vec{x}),\Phi(\vec{x}))\Big)\Psi[h_{ij}(\vec{x}),\Phi(\vec{x})]=0. \quad (1)$$

Here $h_{ij}(\vec{x})$ denotes the metric on three-space, h its determinant, $\Phi(\vec{x})$ symbolically stands for all matter fields, and ${}^{(3)}R$ is the Ricci scalar on three-space (a cosmological constant term has been neglected).

The Wheeler-DeWitt equation (1) does not yield any unification of gravity and matter. One would thus expect that it describes an effective theory which is valid below some unification scale (e.g. of string theory), in the same way as, due to the triviality of the ϕ^4-theory, the standard electroweak theory is considered to be an effective theory which is valid below a certain cut-off given by the Higgs mass. It is thus one of the premises of the canonical approach that (1) is valid at a certain scale between the domain of the "theory of everything" and the domain where classical general relativity is applicable. Moreover, (1) can still be a meaningful theory of "quantum gravity".

Solutions to the Wheeler-DeWitt equation in its full functional form are rare. In the geometrodynamical version (1), this is only possible in the limit of $G \to \infty$, the strong gravity limit, where the term containing the Ricci scalar is absent [1]. The situation improves if one performs a canonical transformation at the classical level, leading to what is known as Ashtekar's variables [2]. Then explicit solutions to the quantum constraints can be given [3]. In contrast to the strong gravity limit, where different space points decouple, the wave functionals now have support on one dimensional loops. This at least demonstrates that a sensible meaning *can* be attributed to quantum general relativity at a non-perturbative level, although the theory is perturbatively nonrenormalisable. It is, however, difficult to interpret the results in this "loop space representation" in terms of the old, more familiar, geometrodynamical language.

In view of this situation, most work has dealt with minisuperspace versions of (1), where all degrees of freedom, except very few like the three scale factors in a Bianchi model, are frozen out. In this context, the main concern has been to study the implementation of certain boundary conditions (see, e.g., [4] and [5]), the construction of a viable Hilbert space structure, and more conceptual issues like the role of time in quantum gravity (see, e.g., [6], [7], [8] and [9]).

One of the important goals in the canonical quantum theory of gravity is of course the establishment of a connection to known physics. The first thing to try is to derive from (1), through a certain approximation scheme, the limit of quantum field theory in a *fixed* classical gravitational background. This has been first investigated in [10] and then studied from various aspects by various authors (see, e.g., [11], [12], [13], [14], [15]). Basically, what one does is to perform an expansion of (1) with respect to the Planck mass, through which one finds the functional Schrödinger equation for matter fields propagating on a classical gravitational background. (Of course, while quantum field theory is recovered here in its Schrödinger picture form, it is assumed that this form is equivalent to doing quantum field theory in the Heisenberg picture, in a fixed classical background).

From this result one may conclude that among the two key ingredients of ordinary quantum field theory, commutation relations and equation of motion, the former is more fundamental. The latter holds true only in the approximation that gravity is classical, and is otherwise to be replaced by the Wheeler-DeWitt equation. The above derivation of quantum field theory from quantum gravity also solves, at least at the mathematical level, the puzzle raised by Smolin [16] regarding the relation between 'quantum and gravitational phenomena'. In the stochastic interpretation of quantum mechanics [17] one assumes the particle's quantum mechanical motion to be a special kind of Brownian motion. The Brownian diffusion constant D is assumed to be proportional to Planck's constant, so that $\hbar = Dm_q$. The constant m_q has dimensions of mass and can vary from particle to particle. Smolin emphasises the coincidence that m_q is "found experimentally to be equal to the particle's inertial mass, to at least 2 parts in 10^{12}". If we accept that non-relativistic quantum mechanics is an approximation to quantum field theory, and that both general relativity and the functional Schrödinger equation are approximations to the Wheeler-DeWitt equation, this coincidence can be understood. The inertial mass (=gravitational mass) and the quantum mass m_q then both have the same origin: namely, the matter stress-energy operator in the Wheeler-DeWitt equation.

If the Schrödinger equation can be *derived* from the Wheeler–DeWitt equation, a natural question to ask is how quantum gravitational corrections to the Schrödinger equation may look like. A first step in this direction has been taken in [18] within the context of a two dimensional minisuperspace model. Here we present such a derivation of these correction terms for the *full* functional equation (1). Although these corrections prove, of course, to be extremely tiny in the laboratory, they show how, at least in principle, effects of quantum gravity

show up through, e.g., a shift in atomic spectral lines.

The paper is organised as follows. In section 2, we present a simple but helpful analogy – an expansion of the Klein–Gordon equation with respect to the speed of light which is formally similar to the case of interest here. In section 3 we define our approximation scheme and derive a functional Schrödinger equation for matter fields on a fixed background. Section 4 comprises the main part of our paper. Going one order further in our approximation scheme, we derive corrections to this functional Schrödinger equation. Furthermore, we apply this scheme to the momentum constraints. We then demonstrate how the corrections show up in the spectrum of hydrogen type atoms. Finally, section 5 contains a brief summary and a critical discussion of the obtained results.

2 A simple analogy: Relativistic corrections to the Schrödinger equation from the Klein–Gordon equation

Consider the Klein–Gordon equation for a (first quantised) wave function $\varphi(\vec{x},t)$

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2\nabla^2 + m^2c^2\right)\varphi(\vec{x},t) = 0. \tag{2}$$

If one compares this with the Wheeler–DeWitt equation (1), one recognises that, formally, there is a correspondence of c^2 in (2) with $c^2/16\pi G$ in (1). The gravitational degrees of freedom (three–metric) in (1) correspond to Minkowski time in (2), while the matter degrees of freedom in (1) correspond to Minkowski three–space in (2). In the following, the nonrelativistic approximation to (2) will be discussed through a Born–Oppenheimer type of approach with c^2 as the parameter. This makes sense, if the relevant velocities are much smaller than the speed of light, as is the case for, e.g., atomic electrons. In this way we will obtain the well-known relativistic corrections to the Schrödinger equation. While our purpose here is to provide an analogy to the discussion of the Wheeler–DeWitt equation, we emphasise that this kind of approximation scheme is mathematically different from the one discussed in the literature, where a Foldy–Wouthysen transformation is applied (see, e.g., [19]).

To make the discussion more general, we insert into (2) a minimally coupled electromagnetic potential A^μ through the usual substitution $p_\mu \to p_\mu - \frac{e}{c}A_\mu$, but keep only its zero component $A^0 \equiv \phi$. One then has instead of (2)

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2\nabla^2 + m^2c^2 + \frac{2ie\hbar}{c^2}\phi\frac{\partial}{\partial t} - \frac{e^2\phi^2}{c^2} + \frac{ie\hbar}{c^2}\frac{\partial\phi}{\partial t}\right)\varphi(\vec{x},t) = 0. \tag{3}$$

We write

$$\varphi(\vec{x},t) = e^{iS(\vec{x},t)/\hbar} \tag{4}$$

and make the expansion

$$S = c^2 S_0 + S_1 + c^{-2} S_2 + \dots \tag{5}$$

We insert this into (3) and compare equal powers of the expansion parameter c^2. To order c^4 we find

$$(\nabla S_0)^2 = 0. \tag{6}$$

Thus S_0 is a function of t only.

The next order (c^2) yields

$$-\left(\frac{\partial S_0}{\partial t}\right)^2 + m^2 = 0. \tag{7}$$

This is a Hamilton–Jacobi type of equation which leads to real solutions for S_0, if $m^2 \geq 0$ ("no tachyons"). It can, of course, be immediately solved,

$$S_0 = \pm mt + const., \tag{8}$$

so that the wave function up to this order reads

$$\varphi \propto e^{\pm imc^2 t/\hbar}. \tag{9}$$

These are the usual wave functions for a particle of positive (lower sign) and negative (upper sign) energy at rest. Note that, in spite of the real nature of (2), these solutions are intrinsically complex. In the following, we will restrict to the positive energy case (lower sign). Of course, it is implicitly assumed throughout that positive and negative energies can be treated separately, i.e. that particle production etc. does not occur.

The next order (c^0) in (3) then yields

$$2m\dot{S}_1 + (\nabla S_1)^2 - i\hbar\nabla^2 S_1 + 2em\phi = 0, \tag{10}$$

which upon introducing $f \equiv e^{iS_1/\hbar}$ can be written as (for $m \neq 0$)

$$i\hbar\dot{f} = -\frac{\hbar^2}{2m}\nabla^2 f + e\phi f. \tag{11}$$

Thus we have recovered nothing but the Schrödinger equation with an external electrostatic potential. The next order in our expansion will therefore yield the first relativistic corrections.

To order c^{-2} we find

$$\begin{aligned} &i\hbar\ddot{S}_1 - i\hbar\nabla^2 S_2 - \dot{S}_1^{\,2} + 2m\dot{S}_2 + 2\nabla S_1 \nabla S_2 \\ &+ie\hbar\dot{\phi} - e^2\phi^2 - 2e\phi\dot{S}_1 = 0. \end{aligned} \tag{12}$$

In this equation we rewrite S_1 in terms of f using the definition given above. Next, we eliminate S_2 by defining $\chi \equiv f e^{iS_2/\hbar c^2}$. Equation (12) can now be interpreted as the Schrödinger equation for the modified wave function χ, and after using (11) for f, (12) can be written as

$$\begin{aligned} i\hbar\dot{\chi} \;=\; &\left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi\right)\chi \\ &+\left(-\frac{\hbar^4}{8m^3c^2}\nabla^2(\nabla^2) + \frac{e\hbar^2}{4m^2c^2}\nabla^2\phi + \frac{e\hbar^2}{2m^2c^2}\nabla\phi\nabla\right)\chi \end{aligned} \tag{13}$$

In this derivation, terms with powers higher than c^{-2} have been neglected for consistency. The result (13) is independent of the value of $\dot{\phi}$.

The terms in the second line of (13) require some explanation. The first term can be understood to arise from expanding the relativistic expression $\sqrt{p^2c^2 + m^2c^4}$ for the energy in powers of p/mc up to order p^4. The second term has the form of a Darwin term (which describes the zitterbewegung), while the third one would correspond – in the case of the Dirac equation – to spin orbit coupling. Here, however, these terms are artifacts of the kind of approximation scheme we have used. First, the expectation value of these terms with respect to any stationary state vanishes. This can be seen as follows: Let $|\psi\rangle$ be a stationary state. Because the Hamiltonian is invariant under time reversal, it can be chosen to be real without loss of generality. The two last terms in (13) are proportional to $(\nabla^2\phi + 2\nabla\phi\nabla)\chi \equiv R$. Then it follows in the position representation that

$$\begin{aligned} \langle\psi|R|\psi\rangle &= \int d^3x\psi(\nabla^2\phi)\psi + 2\int d^3x\psi\nabla\phi\nabla\psi \\ &= \int d^3x\psi(\nabla^2\phi)\psi - 2\int d^3x(\nabla\psi)(\nabla\phi)\psi - 2\int d^3x\psi(\nabla^2\phi)\psi \\ &= -\int d^3x\psi(\nabla^2\phi)\psi - 2\int d^3x\psi(\nabla\phi)(\nabla\psi) = -\langle\psi|R|\psi\rangle \end{aligned}$$

Thus the expectation value has to vanish. Secondly, the two terms under consideration may be absorbed in this order of approximation through a renormalisation of the wave function according to

$$\chi \to \tilde{\chi} \equiv \chi e^{e\phi/mc^2}$$

The renormalised wave function $\tilde{\chi}$ is the wave function which one obtains by applying a Foldy–Wouthuysen transformation (see, e.g., [19], where it is shown that the first "real" Darwin terms in the Klein–Gordon case arise at order c^{-4}). It obeys the corrected Schrödinger equation with only the first term of the second line of (13).

In hydrogen type atoms (with a potential $\phi = -Ze/r$), the first correction term in (13) yields an energy shift (fine structure) according to

$$\begin{aligned}\Delta E_{rel} &= -\int d^3x\psi^*_{nlm}\left(\frac{\hbar^4}{8m^3c^2}\nabla^2\nabla^2\right)\psi_{nlm} \\ &= -\frac{mc^2}{2}(Z\alpha)^4\left(\frac{1}{n^3(l+\frac{1}{2})}-\frac{3}{4n^4}\right). \end{aligned} \tag{14}$$

Here ψ_{nlm} denote the unperturbed wave functions with quantum numbers n, l, and m, and α is the fine structure constant. The Klein–Gordon equation (3) can also be solved exactly for this potential (see, e.g., [20]), leading to (14) through an expansion of the exact energy eigenvalues. There is, of course, the usual problem of how to recover the reduced mass in (14) instead of the electron mass, but this can be satisfactorily dealt with (see again [20]). Of course one knows that (14) gives incorrect values for the experimentally observed fine structure (this was the reason for Schrödinger to reject (3) as a candidate for a quantum mechanical wave equation) and that one has to solve the Dirac equation, taking into account the spin of the electron. This aspect, however, is not important for our analogy. We will see in the next section how a similar formal expansion of the Wheeler–DeWitt equation yields quantum gravitational corrections to the Schrödinger equation.

3 Derivation of the Schrödinger equation from quantum gravity

Our starting point is the full functional Wheeler–DeWitt equation (1), where we perform an expansion with respect to the parameter $M \equiv c^2/32\pi G$. This has a dimension of mass per length, so that we can expect this expansion to be sensible if, for a particle, its rest mass divided by its its Compton wavelength is much smaller than M. This is fulfilled for masses much smaller than the Planck mass.

In analogy to the Klein-Gordon case we write the wave functional $\Psi[h_{ij}(\vec{x}), \phi(\vec{x})]$ as

$$\Psi = e^{iS/\hbar} \tag{15}$$

and expand S in the form

$$S = MS_0 + S_1 + M^{-1}S_2 + ... \tag{16}$$

In the following we use a condensed notation, labelling three–metric components h_{ij} by h_a and components of the DeWitt–metric G_{ijkl} by G_{ab}. This is possible

because indices will appear only in this combination. Then the Wheeler–DeWitt equation reads

$$\left(-\frac{\hbar^2}{2M}[G_{ab}\frac{\delta^2}{\delta h_a\delta h_b}+g_a\frac{\delta}{\delta h_a}]+MV(h_{ab})+h_m(h_{ab},\phi)\right)\Psi = 0. \tag{17}$$

Here V stands for $-2c^2\sqrt{h}^{(3)}R$ (the cosmological constant has been neglected) and h_m stands for the hamiltonian of matter fields. We will assume in the following the presence of scalar fields,

$$h_m = -\frac{\hbar^2}{2}\frac{\delta^2}{\delta\phi^2}+u(h,\phi), \tag{18}$$

but the results will be independent of any special form. The linear derivative term in (17) describes some of the ambiguity in the factor ordering of the kinetic term. One could also include in (17) a term proportional to the curvature scalar in configuration space (see, e.g, [21]), but this would not have any effect on the results discussed below. If one were to demand general covariance in configuration space, one would have to choose the covariant Laplace–Beltrami operator in (17). This requires some care in regularising products of distributions at the same point (see, e.g., [22] and [23] for a discussion of this point). In the following we do not care about that and treat all functional derivatives in a formal sense, as if we were dealing with ordinary derivatives. Surprisingly, it will turn out that, at least formally, the corrections are independent of the factor ordering ambiguity in (17).

We now insert the expansion defined by (15) and (16) into (17) and compare expressions with the same order in M. The highest order (M^2) yields

$$\left(\frac{\delta S_0}{\delta\phi}\right)^2 = 0. \tag{19}$$

If many matter fields were present (which is the realistic case), the term in (19) would have to be replaced by a sum of analogous terms. Thus S_0 depends *only* on the three–metric, provided that all matter fields which are present in (17) have positive definite kinetic terms so that one can conclude that every single term in (19) must vanish. Implicit in this reasoning is, of course, also the assumption that S_0 is chosen to be real. This, however, is necessary, because the gravitational wave function should not correspond to a classically forbidden region (see the following).

The next order (M^1) yields the Hamilton–Jacobi equation for gravity alone,

$$\frac{1}{2}G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta S_0}{\delta h_b}+V(h_{ab}) = 0. \tag{20}$$

In a realistic model, there should also be a matter source on the right hand side of this equation. It was shown in [24] how one can understand the emergence of a "back reaction" term $\langle T_{00}\rangle$ in (20). Here, however, we are interested in quantised matter fields and restrict ourselves to the case (20). It is well known that (20) is – together with the principle of constructive interference – equivalent to all ten (vacuum) Einstein field equations [25]. Once S_0 is given, every three-geometry can be integrated to give a full four dimensional solution of the field equations.

The next order (M^0) yields

$$G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta S_1}{\delta h_b} - \frac{i\hbar}{2}\left(G_{ab}\frac{\delta^2 S_0}{\delta h_a \delta h_b} + g_a\frac{\delta S_0}{\delta h_a}\right) + \frac{1}{2}\left(\frac{\delta S_1}{\delta\phi}\right)^2 - \frac{i\hbar}{2}\frac{\delta^2 S_1}{\delta\phi^2} + u(h_a,\phi) = 0. \tag{21}$$

We now define a functional f according to

$$f \equiv D(h)e^{iS_1/\hbar} \tag{22}$$

and, using a condition on D, will derive an equation for f. From (22) we have

$$i\hbar G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta f}{\delta h_b} = \frac{i\hbar}{D}\frac{\delta D}{\delta h_b}G_{ab}\frac{\delta S_0}{\delta h_a}f - G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta S_1}{\delta h_b}f. \tag{23}$$

From (21) we find

$$h_m f = \frac{i\hbar}{2}\left(G_{ab}\frac{\delta^2 S_0}{\delta h_a \delta h_b} + g_a\frac{\delta S_0}{\delta h_a}\right)f - G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta S_1}{\delta h_b}f. \tag{24}$$

We choose $D(h_{ab})$ to satisfy (compare [11])

$$G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta D}{\delta h_b} - \frac{1}{2}\left(G_{ab}\frac{\delta^2 S_0}{\delta h_a \delta h_b} + g_a\frac{\delta S_0}{\delta h_a}\right)D = 0. \tag{25}$$

Thus, D plays the role of a van–Vleck determinant. It depends explicitly on the factor ordering in (17). For a minisuperspace model with one degree of freedom Q and $g = 0$ one has [18]

$$D = \sqrt{\frac{dS_0}{dQ}}.$$

From (23) and (24) it is then easy to see that f satisfies

$$i\hbar G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta f}{\delta h_b} \equiv i\hbar\frac{\delta f}{\delta\tau} = h_m f. \tag{26}$$

This is the functional Schrödinger equation for matter fields propagating on a *fixed* curved background given by (20). As we remarked in the introduction, such an equation has been derived by various authors using various methods.

The time τ in (26) (which is a "many-fingered time") labels the "trajectories" in superspace which run orthogonal to hypersurfaces $S_0 = constant$. It is usually called WKB–time [26] and plays a prominent role in semiclassical gravity.

To this order of approximation, the wave functional of the system is

$$\Psi = \frac{1}{D}e^{iMS_0/\hbar}f. \tag{27}$$

By computing the wave-functional to the next order, we can find corrections to the Schrödinger equation (26).

We want to conclude this section with one remark concerning the derivation of the Schrödinger equation (26). In its derivation, the gravitational field has been assumed to be in a WKB state. Some authors argue that WKB states are a very special, restricted class of states (see, e.g., [27] and [28]). While we agree that one has to consider an additional mechanism to explain the emergence of classical properties (for example decoherence [14], [34]), the WKB method is the simplest mathematical procedure to establish the quantum to classical correspondence. This remark is reinforced by the results of the present work.

4 Corrections to the Schrödinger equation from quantum gravity

The next order in our expansion (16), namely $O(M^{-1})$, yields the following equation for S_2:

$$G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta S_2}{\delta h_b} + \frac{1}{2}G_{ab}\frac{\delta S_1}{\delta h_a}\frac{\delta S_1}{\delta h_b} - \frac{i\hbar}{2}\left(G_{ab}\frac{\delta^2 S_1}{\delta h_a \delta h_b} + g_a\frac{\delta S_1}{\delta h_a}\right) + \frac{\delta S_1}{\delta\phi}\frac{\delta S_2}{\delta\phi} - \frac{i\hbar}{2}\frac{\delta^2 S_2}{\delta\phi^2} = 0. \tag{28}$$

Proceeding as in the Klein–Gordon case, we rewrite S_1 in terms of f, using the definition (22). Equation (28) then becomes

$$G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta S_2}{\delta h_b} - \frac{\hbar^2}{D^2}G_{ab}\frac{\delta D}{\delta h_a}\frac{\delta D}{\delta h_b} + \frac{\hbar^2}{2D}\left(G_{ab}\frac{\delta^2 D}{\delta h_a \delta h_b} + g_a\frac{\delta D}{\delta h_a}\right) = \frac{\hbar^2}{2f}\left(-\frac{2}{D}G_{ab}\frac{\delta f}{\delta h_a}\frac{\delta D}{\delta h_b} + G_{ab}\frac{\delta^2 f}{\delta h_a \delta h_b} + g_a\frac{\delta f}{\delta h_a}\right) + \frac{i\hbar}{f}\frac{\delta S_2}{\delta\phi}\frac{\delta f}{\delta\phi} + \frac{i\hbar}{2}\frac{\delta^2 S_2}{\delta\phi^2}. \tag{29}$$

Next we write S_2 as $S_2 = \sigma_2(h_a) + \eta(\phi, h_a)$, and by inspecting the left hand side of (29), we choose σ_2 to be the solution of the equation

$$G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta\sigma_2}{\delta h_b} - \frac{\hbar^2}{D^2}G_{ab}\frac{\delta D}{\delta h_a}\frac{\delta D}{\delta h_b} + \frac{\hbar^2}{2D}\left(G_{ab}\frac{\delta^2 D}{\delta h_a \delta h_b} + g_a\frac{\delta D}{\delta h_a}\right) = 0. \tag{30}$$

The physical interpretation for this choice is that $\sigma_2[h_{ab}]$ should correspond to the second WKB order for the gravitational part of the wave functional. For a one-dimensional minisuperspace model the equation for σ_2 reads (with $p \equiv dS_0/dQ \equiv S_0'$)

$$\sigma_2' = -\frac{\hbar^2}{4p^2}p'' + \frac{3\hbar^2}{8}\frac{p'^2}{p^3}$$

and thus is in full analogy to the second order WKB corrections for quantum mechanical wave functions (see, e.g., [30]).

Using (29) and the equation (30) for σ_2, it is easy to write down the equation satisfied by the functional η:

$$G_{ab}\frac{\delta S_0}{\delta h_a}\frac{\delta \eta}{\delta h_b} = \frac{\hbar^2}{2f}\left(-\frac{2}{D}G_{ab}\frac{\delta f}{\delta h_a}\frac{\delta D}{\delta h_b} + G_{ab}\frac{\delta^2 f}{\delta h_a \delta h_b} + g_a\frac{\delta f}{\delta h_a}\right) + \frac{i\hbar}{f}\frac{\delta\eta}{\delta\phi}\frac{\delta f}{\delta\phi} + \frac{i\hbar}{2}\frac{\delta^2\eta}{\delta\phi^2}. \quad (31)$$

The correct wave functional to this order is

$$\Psi(h_a,\phi) = \Psi^{(2)}_{WKB}(h_a) f e^{i\eta/M\hbar}. \quad (32)$$

where

$$\Psi^{(2)}_{WKB} = \frac{1}{D}e^{iMS_0/\hbar}e^{i\sigma_2/M\hbar}. \quad (33)$$

As before we define $\chi = f e^{i\eta/M\hbar}$. This will be the modified Schrödinger wave functional. By using (26) for f and the equation (31) for η we find that χ satisfies the corrected Schrödinger equation

$$i\hbar\frac{\delta\chi}{\delta\tau} = h_m\chi + \frac{\hbar^2}{2Mf}\left(\frac{2}{D}G_{ab}\frac{\delta f}{\delta h_a}\frac{\delta D}{\delta h_b} - G_{ab}\frac{\delta^2 f}{\delta h_a\delta h_b} - g_a\frac{\delta f}{\delta h_a}\right)\chi. \quad (34)$$

This equation is the gravitational analogue of (13). To improve the analogy we now proceed to put (34) in a more transparent form. For this purpose we decompose the derivatives of f with respect to h_a into components normal and tangential to hypersurfaces $S_0 = constant$ in superspace. We write

$$G_{ab}\frac{\delta f}{\delta h_a} = -\frac{i}{\hbar}G_{ab}A^a h_m f + G_{cd}\frac{\delta f}{\delta h_c}l^d l_b. \quad (35)$$

The first term on the right hand side is the component *normal* to $S_0 = constant$. A^a is explicitly given by

$$A^a = \frac{\delta S_0}{\delta h_a}\left(G_{cd}\frac{\delta S_0}{\delta h_c}\frac{\delta S_0}{\delta h_d}\right)^{-1} = -\frac{1}{2V}\frac{\delta S_0}{\delta h_a}, \quad (36)$$

and use has been made of the Schrödinger equation (26) for f. In (36), we have used the Hamilton–Jacobi equation (20).

The second term on the right hand side of (35) is the *tangential* component and l_a denotes a unit vector tangential to $S_0 = const.$, obeying $l^a A_a = 0$, $l^a l_a = 1$. The tangential component is of course not determined by (26). In the following, we will abbreviate this component by a_a.

To decompose the second metric-derivatives in (34) in this way, we first differentiate (35) with respect to h_b and then use (35) to eliminate the first derivatives. This yields

$$G_{ab}\frac{\delta^2 f}{\delta h_a\delta h_b} = -\frac{i}{\hbar}G_{ab}\frac{\delta A^a}{\delta h_b}h_m f - \frac{A_a A^a}{\hbar^2}h_m^2 f - \frac{\delta G_{ab}}{\delta h_b}a^a - \frac{i}{\hbar}A^a h_m a_a + \frac{\delta a_b}{\delta h_b}. \quad (37)$$

From (36) one infers that

$$\frac{\delta A^a}{\delta h_b} = -\frac{1}{2V}\frac{\delta^2 S_0}{\delta h_a\delta h_b} + \frac{1}{2V^2}\frac{\delta S_0}{\delta h_a}\frac{\delta V}{\delta h_b}. \quad (38)$$

Using (35) - (38) and the equation for the van-Vleck determinant (25), one finds for the correction terms of (34) the expression

$$\frac{\hbar^2}{2Mf}\left(-G_{ab}\frac{\delta^2 f}{\delta h_a h_b} + \frac{2}{D}G_{ab}\frac{\delta f}{\delta h_a}\frac{\delta D}{\delta h_b} - g_a\frac{\delta f}{\delta h_a}\right)\chi \equiv B_n + B_t.$$

The "normal" part B_n of the correction reads

$$B_n = -\frac{1}{4MVf}\left(h_m^2 f - \frac{i\hbar}{V}G_{ab}\frac{\delta S_0}{\delta h_b}[\frac{\delta u}{\delta h_a} + \frac{\delta V}{\delta h_a}h_m f]\right)\chi = -\frac{1}{4MVf}\left(h_m^2 f - \frac{i\hbar}{V}[\frac{\delta u}{\delta\tau} + \frac{\delta V}{\delta\tau}\frac{h_m f}{f}]\right)\chi. \quad (39)$$

Because of the use of (25) in this derivation, *all* factor ordering ambiguities have been cancelled in this component.

The "tangential" part B_t reads

$$B_t = -\frac{\hbar^2}{4MVf}\Big(\frac{2}{D}\frac{\delta D}{\delta h_b}[G_{cd}\frac{\delta f}{\delta h_c}l^d]l_b + \frac{\delta G_{ab}}{\delta h_b}[G_{cd}\frac{\delta f}{\delta h_c}l^d]l^a - \frac{\delta}{\delta h_b}[G_{cd}\frac{\delta f}{\delta h_c}l^d]l_b - g_a G_{cd}\frac{\delta f}{\delta h_c}l^d l^a\Big)\chi. \quad (40)$$

Note that

$$\frac{\delta\chi}{\delta h_a} = \frac{1}{f}\frac{\delta f}{\delta h_a}\chi + O(\frac{1}{M}) \quad etc., \quad (41)$$

so that all f-derivatives can be replaced by χ-derivatives in the order under consideration.

It would be worthwhile to comment on the nature of the tangential components in some detail. Firstly, note that in a minisuperspace model with only one gravitational degree of freedom, the derivative of f with respect to h_a is automatically normal to the surface S_0 =constant and there are no tangential components. Secondly, the Schrödinger equation for f does not determine the dependence of f on combinations of the three-metric other than WKB time. Thus we can make the following simplifying assumption. In the functional Schrödinger equation, f is defined as a functional of the matter fields, on a space-like hypersurface labelled by τ. So it appears reasonable to assume that f depends on h_a only through τ. In this case, the "tangential" component is identically zero. (One could, however, also write the terms (40) together with the derivative with respect to τ formally as a derivative of χ with respect to a renormalised WKB time but we will not do this here.)

Thus our final result is the corrected Schrödinger equation

$$i\hbar\frac{\delta\chi}{\delta\tau} = h_m\chi + \frac{4\pi G}{c^4\sqrt{h}^{(3)}R}\left(h_m^2 + \frac{i\hbar}{2\sqrt{h}^{(3)}R}\left[\frac{\delta u}{c^2\delta\tau} - \frac{\delta(2\sqrt{h}^{(3)}R)}{\delta\tau}h_m\right]\right)\chi. \quad (42)$$

This equation is the main result of our paper. We emphasise again that the correction term in (42) is *independent* of the factor ordering which was chosen for the gravitational part of the Wheeler–DeWitt equation – a satisfactory feature. Note also that the correction terms vanish in the nonrelativistic limit $c \to \infty$.

The purpose of discussing the Klein-Gordon example in this paper should now become clear upon comparison of (13) and (42). In both cases one obtains a well-known limit at a certain order, and similar corrections at the next order of approximation. The correction term proportional to h_m^2 appears in both cases, and perhaps the term proportional to h_m in (42) is a "zitterbewegung-like" effect.

The second and third term in the bracket of (42) are imaginary and may lead to a gravitationally induced violation of unitarity. One may, however, expect that their contribution – due to the presence of derivatives with respect to the three-metric – is even smaller than the first correction term.

In canonical quantum gravity, the Wheeler–DeWitt equation (1) has to be supplemented by the momentum constraints. In their general form

$$\left(\frac{1}{\sqrt{h}}\frac{\delta\Psi}{\delta h_{ab}}\right)_{|b} = \frac{8\pi G}{c^4}\sqrt{h}T^{0a}\Psi \quad (43)$$

they guarantee that the wave functional depends – apart from matter fields – only on the three-geometry, i.e. it is invariant under coordinate transformations on three-space. Performing an expansion as in (15), (16), one can easily see that the effect of (43) is to guarantee that, at *each step* of the expansion, the corresponding S depends only on the three geometry (apart from matter fields).

According to the original philosophy of geometrodynamics [31], the Wheeler–DeWitt equation (17) and thus the correction equation (42) hold only for *compact* three-space Σ. In Minkowski or Schwarzschild space, which are asymptotically flat, the correction terms would formally diverge because ${}^{(3)}R = 0$ on hypersurfaces t = *constant*. The primary meaning of (42) should therefore lie within the cosmological context where it provides corrections to quantum field theory on a fixed background. The WKB-time τ would then be directly connected to a certain cosmological solution of Einstein's equations. Take for example a homogeneous scalar field on a closed Friedmann background with scale factor a. One can then integrate (42) over the three-sphere. The corrected Schrödinger equation then reads, using $\int d^3x\sqrt{h}^3R = 12\pi^2 a$ and neglecting the less important imaginary terms on the right hand side of (42),

$$i\hbar\frac{\partial\chi}{\partial\tau} = h_m\chi + \frac{G}{3\pi c^4 a}h_m^2\chi. \quad (44)$$

The corrections are thus utterly negligible, except for very small values of the scale factor.

Although (42), as it stands, cannot hold for asymptotically flat spaces, one might expect on purely dimensional grounds that a length scale analogous to the scale factor plays a prominent role. The term $\sqrt{h}^{(3)}R$ in (42) would then have to be replaced by a term proportional to L_c, where L_c is a typical curvature length of the gravitational background. Consider for example a quantum mechanical hamiltonian of the form

$$h_m(q,h_a) = -\frac{\hbar^2}{2m}\nabla^2 + u(q,h_a). \quad (45)$$

Although the above corrections have been derived for a functional Schrödinger equation, one might expect that these corrections also show up in quantum mechanics (remember the calculation of the Lamb shift). Then the dominant part of the correction terms in (42) reads

$$h_m^2 = \frac{\hbar^4}{4m^2}\nabla^2\nabla^2 - \frac{\hbar^2}{2m}\nabla^2 u - \frac{\hbar^2}{m}\nabla u\nabla - \frac{\hbar^2}{2m}u\nabla^2 + u^2. \quad (46)$$

The second and third term in (46) are analogous to the "Darwin-like" terms in (13) and do not contribute to stationary states. Let us consider in analogy to the Klein–Gordon case (13) the fourth-order derivative term in (46). When inserted into (42), this yields a correction term of the order

$$\frac{G\hbar^4}{12\pi c^4 m^2 L_c}\nabla^2\nabla^2.$$

This, in principle, would lead to an energy shift of the spectral lines for the hydrogen atom analogous to (14):

$$\Delta E_{QG} = \int d^3x\psi^*_{nlm}\frac{G\hbar^4}{12\pi c^4 m^2 L_c}\nabla^2\nabla^2\psi_{nlm}$$

$$= \frac{Gm^2}{3\pi L_c}(Z\alpha)^4\left(\frac{1}{n^3(l+\frac{1}{2})} - \frac{3}{4n^4}\right). \qquad (47)$$

Instead of the rest energy of the electron which occured in (14), we find here an expression that is proportional to the gravitational self–energy of a mass m distributed over a scale L_c. For an electron in the Schwarzschild metric of a proton ($Z = 1$), the typical curvature length scale at the distance of the Bohr radius is about $L_c \approx 3.5 \times 10^{15} cm$. Thus the energy shift in (47) would be of the order $\Delta E_{QG} \approx 3.3 \times 10^{-73} eV$ which will of course be forever unobservable. This is even about thirteen orders of magnitude smaller than the corrections to line widths that would arise through the emission of gravitons in linear quantum gravity [32]. Of course even the correction to the energy shift arising from the classical gravitational interaction between proton and electron is much bigger. Perturbing the potential in (45) by $-Gmm_p/r$, one finds an energy shift

$$\Delta E_G = -\frac{Gm^2 m_p e^2}{\hbar^2 n}$$

which for $n = 1$ is about $1.2 \times 10^{-58} eV$. The signature of this energy shift, however, differs from the one in (47). Thus, at least in principle, there is an effect arising from quantum fluctuations of the gravitational field. The importance of these quantum gravitational corrections lies in the conceptual modification they cause to quantum field theory at the Planck scale.

5 Discussion

We have derived in this paper correction terms to the Schrödinger equation which arise from the coupling of quantum gravitational fluctuations to matter fields. This has been achieved through a formal expansion of the Wheeler–DeWitt equation with respect to powers of the Planck mass. The corrected Schrödinger equation is again a linear equation, though in [18] it was wrongly claimed that the corrected equation is nonlinear in the wave function. We could demonstrate that these correction terms are actually independent of the choice of factor ordering for the kinetic term of the gravitational part. We have also discussed how, in principle, these corrections alter the spectral lines of hydrogen type atoms. As expected, the actual line shift turns out to be extremely tiny and unobservable forever. This is not surprising because atoms are bound by electromagnetic forces, where even the effect of classical gravity is smaller by 39 orders of magnitude. The effect of quantum gravity is another 34 orders of magnitude below that. The only known quantum mechanical system where the influence of classical gravity has been successfully tested is that of neutron beams in the gravitational field of the earth where interference effects

are being induced [33]. But even in that case quantum gravitational effects are suppressed by some 34 orders of magnitude. Thus the situation is hopeless, as far as laboratory experiments are concerned.

The only imaginable situations where these corrections could become important are those of the early universe and the final fate of a black hole. For example, our equation (44) directly yields corrections to the Schrödinger equation for higher multipoles on a Friedmann background [12] which physically represent density fluctuations. While in the cosmological context the interpretation of (42) is more or less clear, this is not true for the black hole case where even the role of the Wheeler–DeWitt equation has not been clarified up to now.

We now discuss the theoretical aspects of (42) and start by comparing these corrections with those in the Klein-Gordon case. Note that even though the Klein-Gordon equation is real, we must choose the leading order solution to be one of the two complex plane waves, to have a sensible one-particle interpretation. For a similar reason, (see, e.g. [15]) we should choose the leading order gravitational wave function to be a complex WKB wave-functional, even though the Wheeler-DeWitt equation is real. Such a choice is necessary if we are to recover the picture of field propagation in a classical, rather than in a superposition of classical universes. To explain the unobservability of such superpositions, one has to invoke an additional mechanism like decoherence [14], [34].

The relativistic corrections found in (13) assume a fixed classical electromagnetic field. They have a straightforward physical interpretation of being due to the relativistic mass increase and the smeared Coulomb potential seen by the relativistic particle. An additional correction (the Lamb shift) arises if quantum fluctuations of the electromagnetic field are taken into account. How does one physically interpret the corrections arising in the Wheeler-DeWitt case? Note that unlike the electromagnetic field in the Klein-Gordon case, now the metric is quantised, and this is expressed in the second order gravitational WKB fluctuations in equation (32). It then appears reasonable to think that the corrections in (42) are, in analogy to the Klein-Gordon case, a "gravity-induced mass increase" and a "gravitational zitterbewegung", being caused by a process we do not yet properly understand. One may think that, instead of the Klein-Gordon equation, a better analogy to the Wheeler-DeWitt case would be provided by the functional Schrödinger equation for a quantised scalar field coupled to a quantised electromagnetic field. However, we do not succeed in finding a natural expansion parameter in the latter case.

Because, for the gravitational part of the total wave functional, our expansion is equivalent to a WKB expansion (and thus to a loop expansion), the WKB(2) fluctuations in (32) should correspond to pure graviton graphs of second order around a classical gravitational background. This is not the case for the χ–part of the wave functional which should contain loops of gravitons and "matter particles" to any order (although no "particles" are defined in this context!). For this reason it is not possible to interpret these corrections straightforwardly as a "Lamb shift-like" effect. One might also ask what is the effect of averaging these WKB(2) fluctuations in expectation values of a matter observable with respect to χ, but this is not attempted here.

In addition to the dominant h_m^2– term in (42) there are further correction terms which yield an imaginary contribution and thus may lead to a quantum gravitational *violation of unitarity*. This is not surprising because the Wheeler–DeWitt equation itself does not posess any unitary behaviour with respect to an "intrinsic time" (which is distinguished by its sign in the kinetic term). The physical reason for this is that all variables are dynamical, so that the wave functional can only give a probability amplitude for time but not in time [26]. This is important because it facilitates the normalisation of the wave functional with respect to all variables [34].

The presence of the h_m^2– term in (42) may have a fundamental significance for quantum field theory. One can no longer expand the wave functional of a free theory into a set of harmonic oscillators as is necessary to relate the functional Schrödinger picture to the more commonly used Fock space representation (see, e.g., [35]). Essentially, the nonlinearity of h_m^2 prevents the separation of the wave functional χ into a product of individual oscillator eigenfunctions. This has drastic consequences for the particle concept. In the approximation of quantum field theory in a fixed background, particles can be defined, but the definition is, in general, not unique. If one takes into account the correction terms derived in this paper, the concept of a particle cannot even be defined.

One of the motivations for the present work is to find a gravity-induced smoothing of divergences in flat space quantum field theory. In principle, this effect should already be contained in (42) and the associated averaging implicit in (32). These two equations together should also imply a lower bound to physical length at the Planck scale. We hope to return to these unresolved issues in a future publication.

Acknowledgements

We thank J. Barbour, F. Krahe, A. Sen and S. Sonego for discussions. TPS thanks Professor Abdus Salam, the International Atomic Energy Agency and UNESCO for hospitality at ICTP, Trieste. Work supported by the Swiss National Science Foundation.

References

[1] C. Teitelboim, Phys. Rev. **D25**, 3159 (1982)

[2] A. Ashtekar, *New Perspectives in Canonical Gravity*, Bibliopolis (Napoli 1988)

[3] T. Jacobson and L. Smolin, Nucl. Phys. **B299**, 295 (1988); C. Rovelli and L. Smolin, Nucl. Phys. **B331**, 80 (1990)

[4] J.B. Hartle and S.W. Hawking, Phys. Rev. **D28**, 2960 (1983)

[5] A. Vilenkin, Phys. Rev. **D37**, 888 (1988)

[6] H.D. Zeh, *The Physical Basis of the Direction of Time*, Springer (Heidelberg 1989)

[7] C. Kiefer, Phys. Rev. **D38**, 1761 (1988)

[8] T. Padmanabhan (1990) *A definition of time in Quantum Cosmology*, preprint TIFR-TAP 6/90

[9] J. Greensite (1990) *Ehrenfest's principle in Quantum Gravity*, preprint SFSU-TH-90

[10] V.G. Lapchinsky and V.A. Rubakov, Acta Physica Polonica **B10**, 1041 (1979)

[11] T. Banks, Nucl. Phys. **B249**, 332 (1985)

[12] J.J. Halliwell and S.W. Hawking, Phys. Rev. **D31**, 1777 (1985)

[13] J.B. Hartle, in *Gravitation and Astrophysics*, J.B. Hartle and B. Carter (eds.), D. Reidel (Dordrecht 1987)

[14] C. Kiefer, Class. Quantum Grav. **4**, 1369 (1987)

[15] T.P. Singh and T. Padmanabhan, Ann. Phys. (N.Y.) **196**, 296 (1989)

[16] L. Smolin, Class. and Quantum Gravity **3**, 347 (1986)

[17] E. Nelson Phys. Rev. **150**, 1079 (1966)

[18] T.P. Singh, Class. Quantum Grav. **7**, L149 (1990)

[19] J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill (New York 1964)

[20] C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, McGraw-Hill (New York 1980)

[21] S.W. Hawking and D.N. Page, Nucl. Phys. **B264**, 185 (1986)

[22] T. Christodoulakis and J. Zanelli, Class. Quantum Grav. **4**, 851 (1987)

[23] N.C. Tsamis and R.P. Woodard, Phys. Rev. **D36**, 3641 (1987)

[24] T. Padmanabhan and T.P. Singh, Class. and Quantum Grav. **7**, 411 (1990)

[25] U.H. Gerlach, Phys. Rev. **177**, 1929 (1969)

[26] H.D. Zeh, Phys. Lett. **A126**, 311 (1988)

[27] J. Barbour and L. Smolin, in *Proc. of the Osgood Hill Meeting on Conceptual Problems in Quantum Gravity*, Birkhäuser (Boston 1990)

[28] E. Joos, Phys. Lett. **A116**, 6 (1986)

[29] H.D. Zeh, Phys. Lett. **A116**, 9 (1986)

[30] L.D. Landau and E.M. Lifshitz, *Quantum Mechanics*, Pergamon Press (Oxford 1975), §46

[31] J.A. Wheeler, in *Battelle Rencontres*, C. DeWitt and J.A. Wheeler (eds.), Benjamin (New York 1968)

[32] S. Weinberg, *Gravitation and Cosmology*, John Wiley (New York 1972), p.285

[33] R. Colella, A.W. Overhauser and S.A. Werner, Phys. Rev. Lett. **34**, 1472 (1975)

[34] H.D. Zeh, *Quantum Measurements and Entropy*, to appear in *Complexity, Entropy, and the Physics of Information*, Santa Fé Institute Studies in the Sciences of Complexity, ed. by W. Zurek, vol. IX. Reading, MA: Addison-Wesley (1990); H.-D. Conradi, Heidelberg preprint (1990)

[35] K. Freese, C.T. Hill and M. Mueller, Nucl. Phys. **B255**, 693 (1985)

Stampato in proprio nella tipografia
del Centro Internazionale di Fisica Teorica