The necessity of quantising the gravitational field is still subject to an open debate. In this paper we compare the approach of quantum gravity, with that of a fundamentally semi-classical theory of gravity, in the weak-field non-relativistic limit. We show that, while in the former case the Schrödinger equation stays linear, in the latter case one ends up with the so-called Schrödinger–Newton equation, which involves a nonlinear, non-local gravitational contribution. We further discuss that the Schrödinger–Newton equation does not describe the collapse of the wave-function, although it was initially proposed for exactly this purpose. Together with the standard collapse postulate, fundamentally semi-classical gravity gives rise to superluminal signalling. A consistent fundamentally semi-classical theory of gravity can therefore only be achieved together with a suitable prescription of the wave-function collapse. We further discuss, how collapse models avoid such superluminal signalling and compare the nonlinearities appearing in these models with those in the Schrödinger–Newton equation.

Keywords: Schrödinger–Newton equation, semi-classical gravity, collapse of the wave-function
1. Introduction

The Schrödinger–Newton equation

\[
\hat{\psi}(t, \mathbf{r}) = \left( -\frac{\hbar^2}{2m} \nabla^2 - Gm^2 \int d^3r' \frac{\psi(t, \mathbf{r}')^2}{|\mathbf{r} - \mathbf{r}'|} \right) \psi(t, \mathbf{r}),
\]

has been brought into play by Diósi [1] and Penrose [2–4] to provide a dynamical description of the collapse of the quantum wave-function. It has regained attention in recent times, mainly due to its connection with the question whether gravity should really be quantized [5], and due to its falsifiability in envisaged experiments [6–8].

There has been quite a debate about how the Schrödinger–Newton equation relates to established fundamental principles of physics [9–12]. Here we wish to clarify the motivations underlying the Schrödinger–Newton equation, in particular, in which sense it follows from semi-classical gravity. We will also discuss its present limitations as a fundamental description of physical phenomena.

First of all one should avoid confusion about what is actually meant by the term semi-classical gravity. Most physicists assume that the gravitational field must be quantized in some way or another, and that semi-classical gravity is only an effective theory, which holds in situations where matter must be treated quantum mechanically, but gravity can be treated classically (although it is fundamentally quantum). In this case, of course, we have a pure quantum theory where everything remains linear. From such a theory we would never expect any nonlinear interaction terms in the Schrödinger equation. Under these assumptions, the Schrödinger–Newton equation can be derived within quantum gravity as a mean-field limit—but only as such—whose validity is restricted to the case of large numbers of particles; we will review this in section 2.15.

The point taken by the proponents of equation (1), however, is different. The Schrödinger–Newton equation does follow from a theory in which only matter fields are quantized, while the gravitational field remains classical even at the fundamental level. This is what we will refer to as semi-classical gravity in the following, thereby adopting the notation of [14–16]. One possible candidate for such a theory is a coupling of gravity to matter by means of the semi-classical Einstein equations

\[
R_{\mu\nu} + \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4} \langle \Psi | \hat{T}_{\mu\nu} | \Psi \rangle,
\]

that is by replacing the classical energy–momentum tensor in Einstein’s equations by the expectation value of the corresponding quantum operator in a given quantum state \( \Psi \). This idea has a long history, dating back to the works of Møller [17] and Rosenfeld [18]. It has been commented repeatedly that such a theory would be incompatible with established principles of physics, namely, through the normal ordering and renormalization prescription. They would lead to a mass-renormalization of the theory rather than a potential term in the Schrödinger equation [9, 10].

The term quantum gravity here is restricted to any such theory which, at least in its low energy limit, treats the gravitational field in the linearized Einstein equations as a linear quantum operator.

Apart from this, the Schrödinger–Newton equation also follows uncontroversially from a gravitating classical matter field [13].
physics [19, 20] but these arguments turn out to be inconclusive [15, 16, 21, 22]. We review them in appendix A.

Note that the validity of (2) at the fundamental level requires that the collapse does not violate local energy–momentum conservation, \( \partial^\mu \langle \Psi | \hat{T}_\mu | \Psi \rangle = 0 \). This is certainly not the case for the standard instantaneous collapse in quantum mechanics. Also collapse models [23–28], that have been constructed to date, violate this condition. There is, however, no obvious reason why local energy–momentum conservation must be violated by any measurement prescription. Indeed Wald [29] shows that such a measurement prescription which is consistent with the semi-classical Einstein equations is possible.

At the current state of physics, the honest answer to the question if the gravitational field must be quantized is therefore that we do not know. The final answer can only be given by experiment. In this regard, it is worthwhile noting that the collapse of the wave function—if it is a real phenomenon—can only be explained by a nonlinear, i.e. non-quantum, interaction. Therefore, if gravity is responsible for the collapse, as often suggested in the literature [30–33], it must remain fundamentally classical (or, in any case, non-quantum), but the form of coupling to quantum matter is of course open to debate.

Given that gravity is fundamentally classical, and that equation (2) is a fundamental equation of nature describing gravity’s coupling to matter, and not an effective equation, the Schrödinger–Newton equation follows naturally. We show this in section 2. However, it is wrong to jump to the conclusion that the Schrödinger–Newton equation alone represents a coherent description of physical phenomena at the non-relativistic level. As we explain in section 3, some collapse rule, or collapse dynamics, must be added in order to account for the stochastic outcomes for measurements of superposition states. But even by adding the standard collapse postulate to the Schrödinger–Newton equation there are difficulties. Quantum non-locality (which is implicit in the collapse postulate together with the Schrödinger–Newton dynamics) leads to the possibility of faster-than-light signalling, as we will show in section 4. In section 5 we compare the Schrödinger–Newton equation with the typical equations used in collapse models, concluding that both types of equations are of a very different structure and cannot easily be combined.

2. The Schrödinger–Newton equation from semi-classical gravity

As anticipated, here we take the point of view that at the fundamental level quantum theory is coupled to classical general relativity via the semi-classical Einstein equation (2). To treat the full equation in the framework of quantum fields on a curved space–time can be a difficult endeavour [29, 34]. But in the linearized theory of gravity [35], where the space–time metric is written as

\[ g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \]

the expansion in \( h_{\mu\nu} \) is well-known to yield the gravitational wave equations at leading order. This remains right for the semi-classical equation (2) where one obtains [35]

\[ \Box h_{\mu\nu} = -\frac{16\pi G}{c^4} \left( \langle \Psi | \hat{T}_{\mu\nu} | \Psi \rangle - \frac{1}{2} \eta_{\mu\nu} \langle \Psi | \eta^{\rho\sigma} \hat{T}_{\rho\sigma} | \Psi \rangle \right), \]

imposing the de Donder gauge-condition \( \partial^\mu (h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} \eta^{\rho\sigma} h_{\rho\sigma}) = 0 \). Here \( \Box \) denotes the d’Alembert operator. Note that the energy–momentum tensor at this order of the linear
approximation is that for flat space–time, while in (2) it was still in curved space–time. In the Newtonian limit, where \( \langle \Psi | \hat{T}_{00} | \Psi \rangle \) is large compared to the other nine components of the energy–momentum tensor, equation (4) becomes the Poisson equation

\[
\nabla^2 V = \frac{4\pi G}{c^2} \langle \Psi | \hat{T}_{00} | \Psi \rangle
\]

for the potential \( V = -\frac{c^2}{2} \hat{h}_{00} \). This is simply the usual behaviour of general relativity in the Newtonian limit: space–time curvature becomes a Newtonian potential sourced by the energy-density term of the energy–momentum tensor.

This potential term now contributes to the Hamiltonian of the matter fields, which in turn yields the dynamics in the Schrödinger equation. To be more specific, in the linearized theory of gravity (3), the interaction between gravity and matter is given by the Hamiltonian [36]

\[
H_{\text{int}} = -\frac{1}{2} \int d^3r \ h_{\mu\nu} \ T^{\mu\nu}.
\]

The quantization of the matter fields then provides us with the corresponding operator:

\[
\hat{H}_{\text{int}} = -\frac{1}{2} \int d^3r \ h_{\mu\nu} \ \hat{T}^{\mu\nu}.
\]

It is important to point out the difference to a quantized theory of gravity. In the latter, \( h_{\mu\nu} \) becomes an operator as well, simply by applying the correspondence principle to the perturbation \( h_{\mu\nu} \) of the metric—and thereby treating the classical \( h_{\mu\nu} \) like a field living on flat space–time rather than a property of space–time. In contrast to this, \( h_{\mu\nu} \) here remains fundamentally classical. It is determined by the wave equations (4), which are meant as classical equations of motion.

In the Newtonian limit, where \( \hat{T}_{00} \) is the dominant term of the energy–momentum tensor, the interaction Hamiltonian then becomes

\[
\hat{H}_{\text{int}} = \int d^3r \ V \ \hat{T}_{00} = -G \int d^3r \ d^3r' \frac{\langle \Psi | \hat{\varrho}(r') | \Psi \rangle}{|r - r'|} \hat{\varrho}(r),
\]

where we already integrated equation (5) and used \( \hat{T}_{00} = c^2 \hat{\varrho} \) in the non-relativistic limit. The mass density operator \( \hat{\varrho} \) is simply \( m \hat{\psi}^\dagger \hat{\psi} \) when only one kind of particle is present. Therefore, following the standard procedure [37] we end up with the Schrödinger–Newton equation in Fock space:

\[
i\hbar \partial_t |\Psi\rangle = \left[ \int d^3r \ \hat{\psi}^\dagger(r) \left( -\frac{\hbar^2}{2m} \nabla^2 + \frac{G m^2}{r} \right) \hat{\psi}(r) \right.
\]

\[
- G m^2 \int d^3r \ d^3r' \frac{\langle \Psi | \hat{\psi}^\dagger(r') \hat{\psi}(r') | \Psi \rangle}{|r - r'|} \hat{\psi}^\dagger(r) \hat{\psi}(r) |\Psi\rangle.
\]

In the non-relativistic limit the number of particles is conserved and we can without further assumptions go over to the first-quantized form. For an \( N \)-particle state
\[ |\Psi_N \rangle = \frac{1}{\sqrt{N!}} \left\{ \int \prod_{i=1}^{N} d^3 r_i \right\} \Psi_N (t, r_1, ..., r_N) \hat{\psi}^+(r_1) \cdots \hat{\psi}^+(r_N) |0\rangle, \]  
(10)

where \( \Psi_N (t, r_1, ..., r_N) \) is the \( N \)-particle wave-function, the expectation value is

\[ \langle \Psi_N | \hat{\psi}^+(r) \hat{\psi} (r) |\Psi_N \rangle = \sum_{j=1}^{N} \left\{ \int \prod_{i=1}^{N} d^3 r_i \right\} |\Psi_N (t, r_1, ..., r_j-1, r, r_{j+1}, ..., r_N) |^2. \]  
(11)

Equation (9) with the state (10) inserted therefore yields the \( N \)-particle Schrödinger–Newton equation [1]

\[ i\hbar \partial_t \Psi_N (t; r_1, ..., r_N) = \left\{ -\sum_{i=1}^{N} \frac{G}{2m} \nabla_i^2 - Gm^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ \int \prod_{k=1}^{N} d^3 k_i \right\} \frac{|\Psi_N (t; r_1, ..., r_N) |^2}{|r_i - r_j|} \right\} \Psi_N (t; r_1, ..., r_N), \]  
(12)

and in the one-particle case the Schrödinger–Newton equation (1) follows immediately.

We therefore unavoidably obtain the Schrödinger–Newton equation for non-relativistic quantum matter if the initial assumptions are correct: that gravity is fundamentally classical, and that the semi-classical Einstein equations (2) describe its coupling to matter. In this precise sense, the Schrödinger–Newton equation does follow from fundamental principles. Whether or not these principles and the underlying assumptions are correct, is a different story which, eventually, will be decided by experiments.

2.1. The Schrödinger–Newton equation as a Hartree approximation

It is important to stress that the Schrödinger–Newton equation also appears in a different context, however with a totally different meaning. Assuming that gravity is fundamentally described by a quantum theory in which the metric perturbation \( h_{\mu\nu} \) turns into a linear operator [38], then very similar to what we have in equation (4), in the weak-field limit one gets the linearized equation:

\[ \square \hat{h}_{\mu\nu} = -\frac{16\pi G}{c^4} \left( \hat{T}_{\mu\nu} - \frac{1}{2} \hat{\eta}_{\mu\nu} \hat{T}_{\rho\sigma} \right), \]  
(13)

where \( \hat{h}_{\mu\nu} \) now is a linear quantum operator. In complete analogy to equation (5), one obtains the potential \( \hat{V} \) in the Newtonian limit where:

\[ \nabla^2 \hat{V} = \frac{4\pi G}{c^2} \hat{T}_{00}. \]  
(14)

Note that, contrary to equation (5), \( \hat{V} \) also carries a hat now, i.e. it is considered a quantum operator. Then, the corresponding interaction Hamiltonian reads as:

\[ \hat{H}_{\text{int}} = \int d^3 r \ \hat{V} \ \hat{T}^{00} = -G \int d^3 r \ d^3 r' \ \frac{\delta (r') \delta (r)}{|r - r'|}. \]  
(15)
Note that the interaction term in equation (15) is in the second-quantized formalism, and we can again write down the full Schrödinger equation in Fock space:

\[
i\hbar \partial_t |\Psi\rangle = \left[ \int d^3r \hat{\psi}^\dagger (\mathbf{r}) \left( -\frac{\hbar^2}{2m} \nabla^2 \right) \hat{\psi} (\mathbf{r}) \right.
\]

\[
- G m^2 \int d^3r d^3r' \frac{\hat{\psi}^\dagger (\mathbf{r}) \hat{\psi} (\mathbf{r}') \hat{\psi}^\dagger (\mathbf{r}') \hat{\psi} (\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \left] |\Psi\rangle. \right. \tag{16}
\]

The corresponding \(N\)-body Schrödinger equation in first-quantized formalism is given by

\[
i\hbar \frac{\partial}{\partial t} \Psi_N(t, \mathbf{r}_1, ..., \mathbf{r}_N) = \hat{H}_N \Psi_N(t, \mathbf{r}_1, ..., \mathbf{r}_N), \tag{17a}\]

where

\[
\hat{H}_N = -\sum_{j=1}^{N} \frac{\hbar^2}{2m_j} \nabla_j^2 - G m^2 \sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{17b}\]

The last term is the contribution of \(\hat{H}_{\text{int}}\) given in equation (15). In the last term on the right-hand side of the above equation, the infinite terms with \(i = j\) are treated by standard renormalization and regularization techniques at the second-quantized level and they appear as a mass renormalization.

Then, in the case of many-particle systems for \(N \to \infty\), and assuming that all particles have the same mass \(m\), one can obtain the nonlinear Hartree equation as the mean-field limit of equation (17a):

\[
i\hbar \frac{\partial}{\partial t} \psi(t, \mathbf{r}) = \left( -\frac{\hbar^2}{2m} \nabla^2 - G m^2 \int d^3r' \frac{\left| \psi (t, \mathbf{r}') \right|^2}{|\mathbf{r} - \mathbf{r}'|} \right) \psi (t, \mathbf{r}). \tag{18}\]

Note that the precise mathematical derivation is not a trivial endeavour and involves implementing the quantum Bogoliubov–Born–Green–Kirkwood–Yvon hierarchy in the limit \(N \to \infty\) [39–41].

Formally, equation (18) is the same as equation (1). However, one should keep in mind that equation (18) is derived as the mean-field limit of an \(N\)-body linear Schrödinger equation and it is only an effective description to the zero-th order of the dynamics of the \(N\)-body system. At the level of the full \(N\)-body system the dynamics are still linear, as given in equation (17a). The centre-of-mass wave-function is described by a free Schrödinger equation.

In contrast to this, in semi-classical gravity we have nonlinearity even at the level of the \(N\)-body dynamics, yielding contributions to the centre-of-mass motion [42]. Therefore, while in the relative motion the non-gravitational interactions dominate the dynamics, in the centre-of-mass dynamics gravitation is the only interaction and can therefore lead to observable effects.

2.2. The Schrödinger–Newton equation and the meaning of the wave-function

Considering the expectation value of the energy–momentum tensor as the source of the gravitational field, the wave-function in the Schrödinger–Newton equation coincides with the gravitational mass distribution, whose different parts attract gravitationally. Following this line of reasoning, a particle would not be really point-like. It would be a wave packet, which tends to
spread out in space, according to the usual dynamical term in the Schrödinger equation, and shrink, according to the Schrödinger–Newton term. An equilibrium is reached when the two effects compensate. This equilibrium state is a soliton, which is physically perceived as a particle.

It can be argued [9, 11] that this mass-density interpretation of the wave-function is in contradiction with the usual probabilistic interpretation, according to which \( \psi(\mathbf{r})^2 \) represents the probability density of finding the (point-like) particle in \( \mathbf{r} \), at the end of a position measurement. On this ground, one could tend to dismiss the Schrödinger–Newton equation. In fact, Schrödinger himself, who had a quite similar interpretation of the wave-function in mind, already noted [43] that in this picture a self-interaction of the wave-function seems to be a natural consequence for the equations to be consistent from a field-theoretic point of view. But he also noticed that, for reasons he could not understand, in the case of electrodynamical interactions such nonlinearities in the Schrödinger equation would give totally wrong numbers for known phenomenology, e.g. the hydrogen spectrum. However, in the case of gravity such an argument does not hold, since all effects of the Schrödinger–Newton equation are far below what has been experimentally observed to date.

It is clear that in a theory based on the Schrödinger–Newton equation, the wave-function must bear both roles: of suitably describing the mass density, and of providing the probability distribution of outcomes of measurements. The question arises: can this be consistently achieved? A positive answer is provided by collapse models; see [23–28] and section 5. In these models, the primary role of the wave-function is to describe matter, meaning with it that a particle is not a point-like particle but is no more and no less than what the wave-function says: a wave packet, which tends to be localized thanks to the collapse mechanism. The important point is that when a particle wave packet interacts with a device which measures its position, the collapse dynamics will say that at the end of the measurement the outcomes will be distributed randomly, according to the Born rule [44]. Therefore, the Born rule is not an additional postulate which assigns a probabilistic role to the wave-function. It is a by-product of the dynamics, when applied to what we typically refer to as measurement processes. It is a handy way—and nothing more—to directly calculate the probability of the outcomes of measurements, rather than solving each time the full equations of motion.

Eventually, the same situation should occur for the Schrödinger–Newton equation, when incorporated in a fully consistent theory. In fact, at the end of the day, the goal of this type of research is to explain the collapse of the wave-function (and with it the Born rule) from an underlying physical principle, which in this case is gravity.

3. Wave-function collapse and the Schrödinger–Newton equation

It has been claimed that the Schrödinger–Newton equation provides an explanation for the collapse of the wave-function [1–4]. In this section we will elaborate on this. The nonlinear gravitational interaction implies an attraction among different parts of the wave-function. When the wave-function of the system is given by a single wave packet, this effect amounts to an inhibition of the free-spread of the wave packet, thus resulting in a self-focusing (or say, shrinking) of it for sufficiently high masses [5, 45]. Additionally, in a system which has been prepared in a spatial superposition of two wave packets at different locations, the nonlinear interaction also implies an attraction between those wave packets (see figure 1).
The strength of the nonlinearity in the Schrödinger–Newton equation depends on the size of the system, in particular on its mass\(^6\). Now we comment on why the attraction between different parts of the wave-function does not account for the usual collapse of the wave-function. Accordingly, exactly as in standard Quantum Mechanics, the collapse postulate, including the Born rule, must be supplemented to the Schrödinger–Newton equation in order to provide a full description of experimental situations.

Let us consider an experiment where a particle’s position is measured. Take an initial superposition state for the particle

\[
\psi(r) = \sqrt{2}\left(\psi_1(r) + \psi_2(r)\right),
\]

where \(\psi_1(r)\) and \(\psi_2(r)\) are wave packets well localized around \(r_1\) and \(r_2\), respectively. During the measurement, this state couples with the massive measuring instrument (say, a pointer) as follows:

\[
\Psi'(r, R) = \frac{1}{\sqrt{2}}\left(\psi_1(r)\Phi_1(R) + \psi_2(r)\Phi_2(R)\right),
\]

where \(\Phi_1(R)\) and \(\Phi_2(R)\) are two localized wave-functions of the pointer, centred around \(R_1\) and \(R_2\), respectively\(^7\). The positions \(R = R_{1,2}\) correspond to the particle being around positions \(r_{1,2}\). Since the pointer is a classical system, according to the orthodox interpretation, the wave-function collapses at \(R = R_1\) or \(R = R_2\), revealing in this way the outcome of the measurement. This means that the particle is found half of the times around the position \(r_1\) and half of the times around the position \(r_2\).

According to the Schrödinger–Newton equation without the standard collapse postulate, on the other hand, a superposition state as in (20) implies a gravitational attraction also between the

---

\(^6\) Denoting the size of a (homogeneous, spherical) particle by \(R\), the width of the wave packet by \(\sigma\), its mass by \(m\), and by \(l_p\) and \(m_p\) the Planck length and Planck mass, respectively, one finds that in the case \(R \ll \sigma\) significant deviations from linear Schrödinger dynamics occur if \(m^3 \sigma^3 \geq m_p^3 l_p\), and in the case \(R \gg \sigma\) they occur if \(m^3 \sigma^3/R \geq m_p^3 l_p\) [6, 42].

\(^7\) In practice these wave-functions will have a finite overlap, but this can be neglected for the purpose of the argument we provide here.
spatial wave packets $\Phi_1$ and $\Phi_2$ representing the massive pointer. The wave-function of the pointer would always ‘collapse’ to the average position $(R_1 + R_2)/2$, simply due to the symmetry of the deterministic dynamics and the initial state. Numerical simulations confirm this behaviour of spatial superpositions collapsing to an average position [45]. Such a behaviour is however in obvious contradiction with the standard collapse postulate, as well as with our everyday experience, where the pointer is found with equal probability either at $R = R_1$ or at $R = R_2$, and never in the middle.

Moreover the Schrödinger–Newton equation is deterministic and as such it cannot explain why quantum measurements occur randomly, distributed according to the Born rule. Therefore, the Schrödinger–Newton equation explains neither the standard collapse postulate nor the Born rule; one still needs both to describe experimental results, as long as no additional collapse prescription is added.

4. Superluminal effects in the Schrödinger–Newton equation

As we pointed out in the introduction, semi-classical gravity together with the standard collapse postulate leads to violation of local energy–momentum conservation. But even if we take the Schrödinger–Newton equation as a hypothesis, without relating it to semi-classical gravity, together with the standard collapse postulate, it leads to superluminal effects, as all nonlinear deterministic Schrödinger equations do [46]. In this section, we discuss a concrete thought experiment, that shows how the Schrödinger–Newton equation implies faster-than-light signalling.

Consider a spin 1/2 particle in a Stern–Gerlach apparatus with a magnetic force in the $z$-direction, where the position of the particle along the $z$-axis is finally observed at the detector. We denote the spin eigenstates along the $z$-axis by $|\pm \rangle_z$ and along the $x$-axis by $|\pm \rangle_x = (|\pm \rangle_z \pm | \mp \rangle_z)/\sqrt{2}$. We assume that there is no coupling of the spin to the spatial wave-function due to gravity, implying that the spatial wave-function evolves according to the Schrödinger–Newton equation while the spin wave-function evolves according to a linear Schrödinger equation.

For the initial spin state $|\pm \rangle_z$ the state after the particle passing through the Stern–Gerlach apparatus is $\Psi(t, R) = \psi_\pm(t, R) \otimes |\pm \rangle_z$. The spatial state evolves as

$$i\hbar \frac{\partial}{\partial t} \psi_\pm(t, R) = \left( -\frac{\hbar^2}{2m} \nabla^2 - Gm^2 \int dz' \frac{|\psi_\mp(t, R')|^2}{|R - R'|} \right) \psi_\pm(t, R), \quad (21)$$

with the initial conditions $\psi_\pm(0, R) = e^{\pm ik_z z} \phi(R)$, where $k_z$ is the momentum induced by the Stern–Gerlach field and $\phi$ is the initial wave packet. Without loss of generality we assume that $\phi$ is a stationary solution of the Schrödinger–Newton equation and therefore the evolution before passing the apparatus plays no role. After passing, the nonlinear gravitational interaction leads to a self-focusing that inhibits the free spreading of the wave packet. The centre of mass moves upwards if the initial state is $\psi_+$ or downwards if it is $\psi_-$. The wave packets $\psi_\pm(t, z)$ are finally observed at the detector positions $\pm d$ (see figure 2(a)).

Now consider the opposite case with initial states $|\pm \rangle_x$. Then the state of the particle after passing through the Stern–Gerlach apparatus is
In this case, the time evolution for the states $\chi^+$ and $\chi^-$ as predicted by the Schrödinger–Newton equation is given by

$$\Psi(t, \mathbf{r}) = \frac{1}{\sqrt{2}} \left( \chi^+(t, \mathbf{r}) \otimes \ket{z^+} + \chi^-(t, \mathbf{r}) \otimes \ket{z^-} \right).$$ (22)

with initial conditions $\chi^\pm(0, \mathbf{r}) = e^{\pm ikz} \phi(\mathbf{r})$ as before. The second term on the right-hand side of equations (23) is the self-focusing force, while the third term corresponds to the attraction between the two wave packets $\chi^+$ and $\chi^-$. Due to this attraction, the wave packets are finally observed at the detector positions $\pm d$, where $d' < d$ (see figure 2(b)). Accordingly, the Schrödinger–Newton equation gives different predictions, compared to the linear Schrödinger equation, for a Stern–Gerlach z-spin measurement with initial states $\ket{z^\pm}$, while it coincides with the linear Schrödinger equation for initial states $\ket{z^\pm}$ if one neglects the change in the width of wave packets for the time-scale of the experiment.

The aforementioned gravitational attraction can, at least in principle, be exploited experimentally to distinguish standard quantum theory from the Schrödinger–Newton equation. For this purpose, we provide a closer look at the conceptual implications of this effect in connection with entanglement and the non-local nature of the quantum measurement process. This effect, that makes use of the nonlinear Schrödinger–Newton dynamics, can then, in principle, be used in order to send signals faster than light.

For that purpose consider a typical EPR set-up, where two particles move in opposite directions toward two Stern–Gerlach devices (see figure 3). They are initially prepared in a singlet spin state:

$$\Psi(t = 0, \mathbf{r}_A, \mathbf{r}_B) = \frac{1}{\sqrt{2}} \left( \ket{z^+}_A \otimes \ket{z^-}_B - \ket{z^-}_A \otimes \ket{z^+}_B \right) \otimes \phi_1(\mathbf{r}_A) \otimes \phi_2(\mathbf{r}_B),$$ (24)
consider an observer (Alice) on the left-hand side of our experimental Stern–Gerlach setting. Alice performs spin measurements either in $z$- or $x$-direction, always before the other entangled particle enters the Stern–Gerlach apparatus on the opposite side, where a second observer (Bob) also is making spin measurements. According to the discussion of the previous section, the measurement is described by the standard collapse rule. Bob always measures the spin in the $z$-direction. If Alice measures the spin in the $z$-direction, then Bob’s particle is prepared in one of the states $\ket{\pm_z}$. Therefore, Bob will detect the particle at positions $\pm d$ (see Figure 3). On the other hand, when Alice measures the spin in the $x$-direction, Bob’s particle will be prepared in the state $\ket{\pm_x}$, meaning that Bob will detect the particle at positions $\pm d’$. Since the distance between Alice and Bob can be arbitrary large, this setup implies faster-than-light signalling.

An important remark is at order. The kind of faster-than-light signalling discussed in this section is an effect of the instantaneous collapse of the wave-function (as a result of

---

Figure 3. Two entangled particles are sent from the source (S) to Alice (A) and Bob (B). Alice is the first one who performs the spin measurement, which can be done along the $z$ or $x$ direction. In the first case, the particle on Bob’s side has a well defined spin along the $z$ direction, and Bob will find the particle around either points $\pm d$. On the contrary, in the second case the particle on Bob’s side is in a superposition of states with well defined spin along the $z$, which implies that Bob will find the particle around either points $\pm d’$. Although, in general, due to the weakness of the gravitational interaction, the difference between $d$ and $d’$ will be incredibly hard to measure, in principle, Bob can identify which measurement Alice performed by only looking at the position where he observes the particle hitting the screen. Therefore the Schrödinger–Newton equation allows to send signals faster than the speed of light.

---

To give some numbers, consider a particle travelling with velocity $v \ll c$ over distance $S$ to the Stern–Gerlach device, where it is split in a superposition with spatial separation $d_0$. The two states in the superposition then travel parallel to each other over a distance $s$. Considering the Newtonian gravitational acceleration one obtains $\Delta d \approx \frac{Gm s^2}{(2v^2d_0^2)}$ as an estimate for the distance they are shifted. For the communication to be superluminal the time it would take for a photon to reach the Stern–Gerlach device, $S/c$, must be larger than than the time $s/v$ it takes for the particle to acquire a measurable displacement $\Delta d$. Therefore $S > c d_0 \sqrt{2\Delta d/(Gm)}$ must hold. Current experiments can achieve $m \approx 10 000 \mu$ and $d_0 \approx \Delta d \approx 1 \mu$m, yielding a minimum distance of about one light-year for superluminal signalling with state-of-the-art technology.
Alice’s measurement), together with the nonlinear character of the dynamics described by the Schrödinger–Newton equation. Therefore, even if one describes the whole situation in a fully relativistic way (i.e. by some sort of ‘Dirac–Newton equation’, which one could eventually obtain by applying (2) to a Dirac field), one would not get rid of the instantaneous collapse of the wave-function upon measurement, nor of the nonlinear character of the dynamics. What would change is the way the two parts of the superposition attract each other: in the Schrödinger–Newton equation this attraction is instantaneous, while in the relativistic framework it would likely have a finite speed. This amounts in slight differences in the self-gravitation effects, which do not play any important role for the argument proposed here. As long as there is some measurable effect of self-gravitation, Bob can always exploit it to figure out Alice’s measurement setting, and thereby receive a signal with the ‘speed of collapse’ (which is infinite in the standard collapse prescription and has been shown to exceed the speed of light by orders of magnitude in a multitude of experiments [47–49]).

Contrary to this situation, it has been widely studied how to modify Schrödinger equation by adding nonlinear and stochastic terms, in order to describe the collapse of the wave-function, while avoiding superluminality. Collapse models provide a mathematically consistent phenomenology of this type [23–28]. In the next section, we elaborate on this issue and discuss the connection of the Schrödinger–Newton equation with collapse models.

5. Comparison between the Schrödinger–Newton equation and collapse equations

The reason why no superluminal effects appear in collapse models [23–28] is that the nonlinear modification is balanced by appropriate stochastic contributions. At the density-matrix level the collapse dynamics are generally of Lindblad type:

$$\frac{d}{dt} \rho_t = -\frac{i}{\hbar} [\hat{H}, \rho_t] + \gamma \int d^3 k \left( \hat{L}(k) \rho_t \hat{L}(k) \right)^\dagger - \frac{1}{2} \hat{L}(k) \hat{L}(k) \rho_t - \frac{1}{2} \rho_t \hat{L}(k) \hat{L}(k),$$

(25)

where both the linear operator $\hat{L}$ and the real coupling constant $\gamma$ can be chosen arbitrarily and must be specified in concrete models. As one cause, at the density-matrix level, stochastic terms perfectly cancel all nonlinear terms.

At the wave-function level, on the other hand, both nonlinear and stochastic contributions appear, and the dynamical equation is of the following type:

$$\frac{\partial}{\partial t} |\psi_t\rangle = \left[ -\frac{i}{\hbar} \hat{H} + \sqrt{\gamma} \int d^3 k \left( \hat{L}(k) - \ell_t(k) \right) \xi(t, k) \right.\left. \right. - \frac{\gamma}{2} \int d^3 k \left( \hat{L}(k) \hat{L}(k) - 2 \ell_t(k) \hat{L}(k) + \left| \ell_t(k) \right|^2 \right) |\psi_t\rangle.$$  

(26)

In this modified Schrödinger equation, nonlinearities are introduced by the expectation-value term

$$\ell_t(k) \equiv \frac{1}{2} \left( \langle \hat{L}(k) \rangle_t + \langle \hat{L}^\dagger(k) \rangle_t \right); \ \ \ \text{with} \ \ \ \langle \hat{L}(k) \rangle_t = \langle \psi_t | \hat{L}(k) | \psi_t \rangle$$

(27)
coupled to a random noise $\xi(t, \mathbf{k}) = dW(t, \mathbf{k})/dt$, where $W(t, \mathbf{k})$ are independent Wiener processes.

How does this compare to the Schrödinger–Newton equation? In order to see this, first note that the Fourier transform of the Newtonian gravitational potential term is given by \[ \frac{G m^2}{|\mathbf{r} - \mathbf{r}'|} = -\frac{G m^2}{2\pi^2} \int d^3k \frac{\exp(i \mathbf{k} \cdot (\mathbf{r} - \mathbf{r}'))}{k^2}. \] (28)

If we now introduce the linear operator
\[ \hat{L}(\mathbf{k}) = m \frac{\exp(i \mathbf{k} \cdot \hat{\mathbf{r}})}{k}, \] (29)
we can write the Schrödinger–Newton equation in terms of solely this operator and its expectation value:
\[ \frac{\partial}{\partial t} |\psi_t\rangle = \left( -\frac{i}{\hbar} \hat{H} + i \frac{G}{2\pi^2\hbar} \int d^3k \left\langle \hat{L}^\dagger(\mathbf{k}) \right\rangle_t \hat{L}(\mathbf{k}) \right) |\psi_t\rangle. \] (30)

As we see, the equation is of a completely different structure with respect to the collapse equation (26). In particular, the coupling constant in front of the nonlinear term is imaginary while $\gamma$ in equation (26) is real.

We could obtain a term as in equation (30), if in equation (26) we replace the noise field by
\[ \xi(t, \mathbf{k}) \rightarrow \xi'(t, \mathbf{k}) + i \left\langle \hat{L}^\dagger(\mathbf{k}) \right\rangle_t, \] (31)
thereby introducing an imaginary drift. The Schrödinger–Newton term would however appear with many other terms which completely change the dynamics.

The collapse equation that comes closest to combining equation (26) with gravity is the collapse model by Diósi [30–32]. This model, in its original form, is obtained by using the operator (29) and the constant $\gamma = G/(2\pi^2\hbar)$ in equation (26).

This equation, however, yields an infinite rate for the energy exchange (i.e., $\text{Tr} (\hat{H} \frac{d\hat{\rho}}{dt})$ diverges). To avoid this problem, one needs to modify the Lindblad operator, introducing a cut-off $R_0$:
\[ \hat{L}(\mathbf{k}) = m \rho(k, R_0) \exp(i \mathbf{k} \cdot \hat{\mathbf{r}}) \frac{\exp(i \mathbf{k} \cdot \hat{\mathbf{r}})}{k}. \] (32)

A reasonable choice for the cut-off is $\rho(k, R_0) \sim e^{-k^2R_0^2}$. The value of $R_0$ was originally chosen equal to the classical size of a nucleon ($\sim 10^{-15}$ m) [30, 32]. Later on, it was shown [51] that such a cut-off is too small, and still giving rise to an unacceptable energy increase of $\sim 10^{-4}$ K s$^{-1}$ for a proton. This problem can be fixed by choosing a much larger cut-off, e.g. $R_0 \sim 10^{-7}$ m, bringing the energy increase down to about $10^{-28}$ K s$^{-1}$.

Although such a model provides a consistent description of the wave-function collapse which does not allow for faster-than-light signalling, its relation to gravity is restricted to the appearance of the gravitational constant $G$ in the coupling. It is an effective model which is not derived from known fundamental principles of physics and it is not clear how this could be done. In fact, the dynamical equation (26) with the operator (32) is simply postulated. In
particular, the newly introduced free parameter $R_0$ is not related to gravity, and the origin of the stochastic term $\xi$ remains unresolved.

6. Conclusions

The main goal of this paper was to straighten out the conceptual status of the Schrödinger–Newton equation, and thereby also clarifying the statements already made in [1, 9, 10, 13, 42]. We have shown that the Schrödinger–Newton equation follows without further assumptions from a semi-classical theory of gravity, i.e. a theory which treats gravity as fundamentally classical with quantum matter coupled to it via the semi-classical Einstein equations (2). The evolution according to the Schrödinger–Newton equation differs from the linear Schrödinger equation in two respects. First, the Schrödinger–Newton dynamics lead to a self-focusing of wave packet solutions for the centre of mass, as it has been studied in [5–7]. Second, in contradiction to the probability interpretation of the wave-function, a spatial superposition state will reveal a Newtonian gravitational attraction of different parts of the wave-function.

The Schrödinger–Newton equation does, however, not serve the purpose it was originally considered for, namely of providing an explanation for the wave-function collapse [3]. A collapse prescription, either in terms of the Copenhagen collapse postulate or in terms of an objective collapse model, is still necessary to relate solutions of the Schrödinger–Newton equation to outcomes of measurements. Moreover, we explicitly have shown that with the conventional collapse prescription, the Schrödinger–Newton dynamics unavoidably lead to the possibility of faster-than-light signalling.

A possible way to describe the collapse of the wave-function without having faster-than-light signalling is given by collapse models. These models avoid these superluminal effects because of the random nature of the collapse dynamics. Therefore the question arises, if the gravitational self-interaction of the wave-function according to the Schrödinger–Newton equation can be brought together with the collapse dynamics. But the way collapse models introduce nonlinearities is very different from that of the Schrödinger–Newton equation. Also, the already known collapse models that are inspired by gravity [30–32] do not fulfil the purpose of combining both ideas since they do not give a satisfactory explanation of how the interactions that lead to the collapse derive from gravity.

Even if the Newtonian gravitational interaction could be consistently included in collapse models, the presence of both Schrödinger–Newton and collapse terms would not resolve the main open question regarding collapse models, namely the physical nature of the stochastic field causing the collapse. If one wants to attribute the collapse to gravity, the gravitational interaction has to account for random effects, either instead of or on top of the semi-classical theory.

Acknowledgements

We would like to thank the anonymous referees for helpful comments. We gratefully acknowledge funding and support through the EU project NANOQUESTFIT, the COST Action MP1006 ‘Fundamental Problems in Quantum Physics’, the John Templeton foundation (grant 39530), and the Istituto Nazionale di Fisica Nucleare (INFN).
Appendix A. Can gravity be fundamentally classical?

The question of the validity of the Schrödinger–Newton equation as a correction to the linear one-particle Schrödinger equation is directly related to the question if there is any need for a quantization of the gravitational field. The great success of quantum theory led to the prevalent belief that such a quantum theory of gravity must be found, by modifying general relativity in order to make it compatible with linear quantum theory. But there is no a priori argument rendering this approach more valid than the opposite one, i.e. retaining the classical structure of gravity and modifying quantum theory. One should keep in mind that, after all, the gravitational field describes the properties of space–time, namely its curvature. Therefore it somehow differs from the other fields which are living on that space–time. If it is really the right approach to consider space–time curvature as a field living on space–time is debatable. But without this view of gravity as ‘just another field’ there is no reason to quantize gravity at the first place. As Rosenfeld [18] points out, the question if the gravitational field has to be quantized is not for theory to decide but for the experiment.

Contrary to frequent claims, there is no conclusive evidence, neither experimentally nor theoretically, that gravity cannot be fundamentally classical. In fact, a fully consistent model to describe the interaction of a classical gravitational and a quantized field was given by Albers et al [22]. They show that a two-dimensional version of Nordström’s scalar theory of gravity can be coupled to a quantized massive scalar field without giving rise to any inconsistencies. This obviously disproves claims that quantization of all fields follows as a necessary requirement of mathematical consistency of any quantum theory. Therefore, this result is also in clear contradiction with a frequently quoted thought experiment by Eppley and Hannah [19].

Eppley and Hannah claim that any theory that couples classical gravity to quantum matter unavoidably leads to inconsistencies, no matter what nature this coupling is. Their thought experiment is based on the scattering of a classical gravitational wave with a quantum particle and distinguishes between two situations.

In the first situation, they assume that this scattering acts as a quantum measurement and therefore collapses the wave-function. In this scenario, the classical wave—for which the de Broglie relation does not hold—is used to measure the position of a particle of definite momentum, and in this way it violates the uncertainty principle. There are at least three problems with this argument. First of all, the analysis by Albers et al [22] shows that even for a classical-quantum-coupling the classical field inherits part of the uncertainties of the quantum field it is coupled to. Thus, it is not true that the classicality of the gravitational wave allows for an arbitrarily precise position measurement. Second of all, the uncertainty relations are a mere corollary of linear quantum theory rather than a basic ingredient of the theory and are not experimentally tested in the situation at hand. They might readily be violated in this case. And finally, Eppley and Hannah provide a detailed description of their experimental set-up for which it becomes apparent that it is not realizable—not even in principle—within the parameters of our universe. In particular, the experiment would require a tremendously massive set of detectors. They would have to be so massive that the detector arrangement would unavoidably be located within its own Schwarzschild radius [21]. Such a coupling of classical gravity to quantum matter, in which a gravitational measurement collapses the wave-function, therefore yields no obvious contradiction to fundamental principles of physics.

The semi-classical Einstein equations (2) as well as the Schrödinger–Newton equation (1), however, belong to the second situation which Eppley and Hannah consider: that the scattering
of a gravitational wave leaves the wave-function intact. For this scenario they construct a
different type of thought experiment in which the scattering of a gravitational wave is used to
probe the shape of the wave-function (instead of the expectation values of the observables as
usual in quantum mechanics). Making use of this in an EPR-like set-up opens the possibility of
faster-than-light signalling. One could ask again if the experiment can be conducted, at least in
principle. In fact Eppley and Hannah elaborate much less on this second experiment than on
their first one and similar arguments as in the first case might also render this second thought
experiment not feasible. But at the bottom of this is probably the incompatibility of the realistic
interpretation of the wave-function in the semi-classical equations (2) with the instantaneous,
non-local Copenhagen collapse. This gets even more evident in consideration of the thought
experiment allowing superluminous signalling by means of the Schrödinger–Newton
equation which we presented in section 4. It could very well be, for example, that the problem
lies with the collapse of the wave-function, more than with the way gravity is treated.

There is a second experiment by Page and Geilker [20, 52] which is often quoted as an
argument for the necessity of a quantization of the gravitational field. In fact, a very similar idea
to that of Page and Geilker was already pronounced by Kibble [14]. He suggests a black-box in
which a quantum decision-making process is used to create a macroscopic superposition state.
According to semi-classical gravity the gravitational field of such a black-box containing a
superposition state would differ from the gravitational field of one containing a collapsed state.
Page and Geilker, however, claim to actually conduct this experiment and find no evidence for
such a difference of the classical gravitational field. But instead of having a black-box, in their
experiment the so-called ‘superposition state’ is created in a purely classical procedure which is
completely decoupled from the decision-making process. Therefore the superposition exists
only in a no-collapse interpretation of quantum mechanics and only the inconsistency of such an
interpretation with semi-classical gravity is shown. The authors try to dissolve this obvious flaw
in their reasoning by arguing that an instantaneous collapse would contradict the divergence-
free nature of the Einstein equations but this is certainly not a problem for any form of
dynamical description of the collapse. The conclusion therefore remains what Kibble had
already noticed: that this is an indication that the connection of gravity and quantum mechanics
requires understanding the wave-function collapse.

References

[1] Diósi L 1984 Phys. Lett. A 105 199–202
[2] Penrose R 1996 Gen. Relativ. Gravit. 28 581–600
[3] Penrose R 1998 Phil. Trans. R. Soc. London 356 1927–39
[4] Penrose R 2014 Found. Phys. 44 557–75
[5] Carlip S 2008 Class. Quantum Grav. 25 154010
[6] Giulini D and Großardt A 2011 Class. Quantum Grav. 28 195026
[7] Giulini D and Großardt A 2013 Class. Quantum Grav. 30 155018
[8] Yang H, Miao H, Lee D S, Helou B and Chen Y 2013 Phys. Rev. Lett. 110 170401
[9] Anastopoulos C and Hu B L 2014 Newton–Schrödinger equations are not derivable from general relativity +
quantum field theory arXiv:1402.3813
[10] Anastopoulos C and Hu B L 2014 Problems with the Newton–Schrödinger equations New J. Phys. 16 085007
[11] Adler S L 2007 J. Phys. A: Math. Theor. 40 755–63
[12] Christian J 1997 Phys. Rev. D 56 4844–77
[13] Giulini D and Großardt A 2012 *Class. Quantum Grav.* **29** 215010
[14] Kibble T W B 1981 Is a semi-classical theory of gravity viable? *Quantum Gravity 2 A Second Oxford Symposium* ed C J Isham, R Penrose and D W Sciama (Oxford: Oxford University Press) pp 63–80
[15] Mattingly J 2005 Is quantum gravity necessary? *The Universe of General Relativity (Einstein Studies* vol 11) ed A J Kox and J Eisenstaedt (Boston: Birkhäuser) pp 327–38 chapter 17
[16] Kiefer C 2007 Quantum gravity 2nd edn *International Series of Monographs on Physics* vol 124 (Oxford: Clarendon)
[17] Møller C 1962 *Les théories relativistes de la gravitation (Colloques Internationaux CNRS* vol 91) ed A L M A Tonnelat (Paris: CNRS) pp 1–96
[18] Rosenfeld L 1963 *Nucl. Phys.* **40** 353–6
[19] Eppley K and Hannah E 1977 *Found. Phys.* **7** 51–68
[20] Page D N and Geilker C D 1981 *Phys. Rev. Lett.* **47** 979–82
[21] Mattingly J 2006 *Phys. Rev. D* **73** 064025
[22] Albers M, Kiefer C and Reginatto M 2008 *Phys. Rev. D* **78** 064051
[23] Ghirardi G C, Rimini A and Weber T 1986 *Phys. Rev. D* **34** 470–91
[24] Ghirardi G C, Pearle P and Rimini A 1990 *Phys. Rev. A* **42** 78–89
[25] Adler S L and Bassi A 2007 *J. Phys. A: Math. Theor.* **40** 15083–98
[26] Adler S L and Bassi A 2009 *Science* **325** 275–6
[27] Bassi A and Ghirardi G 2003 *Phys. Rep.* **379** 257–426
[28] Bassi A, Lochan K, Satin S, Singh T P and Ulbricht H 2013 *Rev. Mod. Phys.* **85** 471–527
[29] Wald R M 1994 *Quantum field theory in curved spacetime and black hole thermodynamics* Chicago Lectures in Physics (Chicago: University of Chicago Press)
[30] Diósi L 1987 *Phys. Lett. A* **120** 377–81
[31] Diósi L 1989 *Phys. Rev. A* **40** 1165–74
[32] Diósi L 2007 *J. Phys. A: Math. Theor.* **40** 2989–95
[33] Adler S L 2014 Gravitation and the noise needed in objective reduction models arXiv:1401.0353
[34] Birrell N D and Davies P C W 1982 *Quantum Fields in Curved Space* (Cambridge: Cambridge University Press)
[35] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman) pp 435–47 chapter 18
[36] Maggiore M 2008 *Gravitational Waves (Theory and Experiments* vol 1) (Oxford: Oxford University Press) p 73 chapter 2
[37] Robertson B 1972 *Am. J. Phys.* **41** 678–90
[38] Donoghue J F 1994 *Phys. Rev. D* **50** 3874–88
[39] Erdős L and Yau H T 2001 *Adv. Theor. Math. Phys.* **5** 1169–205
[40] Bardos C, Erdős L, Golse F, Mauser N and Yau H T 2002 *C. R. Acad. Sci., Paris* **334** 515–20
[41] Benedikter N, Porta M and Schlein B 2013 Mean-field evolution of fermionic systems arXiv:1305.2768
[42] Giulini D and Großardt A 2014 *New J. Phys.* **16** 075005
[43] Schrödinger E 1927 *Ann. Phys. (Berlin)* **82** 265–72
[44] Bassi A and Salvetti D G M 2007 *J. Phys. A: Math. Theor.* **40** 9859–76
[45] Harrison R, Moroz I and Tod K P 2003 *Nonlinearity* **16** 101
[46] Gisin N 1989 *Helv. Phys. Acta* **62** 363–71
[47] Stefanov A, Zbinden H, Gisin N and Suarez A 2002 *Phys. Rev. Lett.* **88** 120404
[48] Scarani V, Bancal J D, Suarez A and Gisin N 2014 *Found. Phys.* **44** 523–31
[49] Bancal J D, Pironio S, Acín A, Liang Y C, Scarani V and Gisin N 2012 *Nat. Phys.* **8** 867–70
[50] Jackson J D 1962 *Classical Electrodynamics* (New York: Wiley) p 88 chapter 3
[51] Ghirardi G, Grassi R and Rimini A 1990 *Phys. Rev. A* **42** 1057–64
[52] Page D N and Geilker C D 1982 An experimental test of quantum gravity Quantum Structure Space Time. *Proc. Nuffield Workshop (Imperial College London 3–21 August 1981)* ed M J Duff and C J Isham (Cambridge: Cambridge University Press) pp 11–21