The Wave Function and Its Evolution

Shan Gao*

September 9, 2011

Abstract

The meaning of the wave function and its evolution are investigated. First, we argue that the wave function in quantum mechanics is a description of random discontinuous motion of particles, and the modulus square of the wave function gives the probability density of the particles being in certain locations in space. Next, we show that the linear non-relativistic evolution of the wave function of an isolated system obeys the free Schrödinger equation due to the requirements of spacetime translation invariance and relativistic invariance. Thirdly, we argue that the random discontinuous motion of particles may lead to a stochastic, nonlinear collapse evolution of the wave function. A discrete model of energy-conserved wavefunction collapse is proposed and shown consistent with existing experiments and our macroscopic experience. Besides, we also give a critical analysis of the de Broglie-Bohm theory, the many-worlds interpretation and other dynamical collapse theories, and briefly discuss the issues of unifying quantum mechanics and relativity.

Contents

*Unit for HPS and Centre for Time, University of Sydney, NSW 2006, Australia. E-mail: sgao7319@uni.sydney.edu.au.

1 Introduction

Quantum mechanics is a physical theory about the wave function and its evolution. There are two main problems in the conceptual foundations of quantum mechanics. The first one concerns the physical meaning of the wave function in the theory. It has been widely argued that Born's probability interpretation is not entirely satisfactory because of resorting to the vague concept - measurement (Bell 1990), though it is still the standard interpretation in textbooks nowadays. On the other hand, the meaning of the wave function is also in dispute in the alternative formulations of quantum mechanics such as the de Broglie-Bohm theory and the many-worlds

interpretation (de Broglie 1928; Bohm 1952; Everett 1957). Exactly what does the wave function describe then?

The second problem concerns the evolution of the wave function. It includes two parts. One part concerns the linear Schrödinger evolution. Why does the linear non-relativistic evolution of the wave function satisfy the Schrödinger equation? A satisfactory derivation of the equation seems still missing (cf. Nelson 1966). The other part concerns the collapse of the wave function during a measurement, which is usually called the measurement problem. It is still unknown whether the wavefunction collapse is real or apparent. Even if the wave function does collapse under some circumstances, it is unclear why and how the wave function collapses either. The measurement problem has been widely acknowledged as one of the hardest and most important problems in the foundations of quantum mechanics.

In this paper, we will try to solve these two fundamental problems of quantum mechanics. In Section 2, we first investigate the physical meaning of the wave function. It is shown that the mass and charge density of a quantum system, which is measurable by protective measurement and proportional to the modulus square of its wave function, is not real but effective, and it is formed by the time average of the ergodic motion of a localized particle with the total mass and charge of the system. Moreover, it is argued that the ergodic motion is not continuous but discontinuous and random. This results then suggests that the wave function in quantum mechanics is a description of random discontinuous motion of particles, and the modulus square of the wave function gives the probability density of the particles *being* in certain locations in real space. In Section 3, we analyze the linear evolution of the wave function. It is shown that the linear non-relativistic evolution of the wave function of an isolated system obeys the free Schrödinger equation due to the requirements of spacetime translation invariance and relativistic invariance. In addition, we also analyze the physical basis and meaning of the principle of conservation of energy and momentum in quantum mechanics.

In Section 4, we investigate the implications of the suggested interpretation of the wave function for the solution to the measurement problem. It is first shown that the two main quantum theories without wavefunction collapse, namely the de Broglie-Bohm theory and the many-worlds interpretation, are inconsistent with the picture of random discontinuous motion of particles. This result suggests that the wavefunction collapse is probably a real physical process. Next, it is argued that the random discontinuous motion of particles may provide an appropriate random source to collapse the wave function. Moreover, the wavefunction collapse is an essentially discrete process due to the discontinuity of motion, and the collapse states are energy eigenstates when the principle of conservation of energy is satisfied. Based on these analyses, we further propose a discrete model of energy-conserved wavefunction collapse. It is shown that the model is consistent

with existing experiments and our macroscopic experience. Besides, we also give some critical comments on other dynamical collapse models, mainly including Penrose's gravity-induced collapse model and the CSL (Continuous Spontaneous Localization) model. Conclusions are given in the last section.

2 Meaning of the wave function

The physical meaning of the wave function is the first interpretative problem of quantum mechanics. Notwithstanding more than eighty years' developments of the theory, however, it is still a hotly debated issue. Besides the standard probability interpretation in textbooks, there are also various conflicting views on the wave function in the alternative formulations of quantum mechanics. In this section, we will try to solve this fundamental interpretive problem through an analysis of the mass and charge density of a single quantum system.

2.1 How do the mass and charge of a quantum system distribute?

The mass and charge of a classical system always localize in a definite position in space at each moment. For a charged quantum system described by the wave function $\psi(x,t)$, how do its mass and charge distribute in space then? We can measure the total mass and charge of the quantum system by gravitational and electromagnetic interactions and find them in certain region of space. Thus it seems that the mass and charge of a quantum system must also exist in space with a certain distribution. In the following, we will first give a heuristic argument for this conjecture and then discuss the more convincing answer given by protective measurement (Aharonov and Vaidman 1993; Aharonov, Anandan and Vaidman 1993; Aharonov, Anandan and Vaidman 1996; Vaidman 2009).

2.1.1 A heuristic argument

The Schrödinger equation of a charged quantum system under an external electromagnetic potential may provide a clue to the answer. The equation is

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=[-\frac{\hbar^2}{2m}(\nabla-\frac{iQ}{\hbar c}A)^2+Q\varphi]\psi(x,t), \tag{1}$$

where m and Q are the mass and charge of the system, respectively, φ and A are the electromagnetic potential, and c is the speed of light. The electrostatic interaction term $Q\varphi\psi(x,t)$ in the equation indicates that the interaction exists in all regions where the wave function of the system, $\psi(x,t)$, is nonzero, and thus it seems to suggest that the charge of the

system also distributes throughout these regions. If the charge does not distribute in some regions where the wave function is nonzero, then there will not exist an electrostatic interaction there. Furthermore, since the integral $\int_{-\infty}^{\infty} Q|\psi(x,t)|^2 d^3x$ is the total charge of the system, the charge density of the system, if indeed exists, will be $Q|\psi(x,t)|^2$. Similarly, the mass density of a quantum system can be obtained from the Schrödinger equation of the system under an external gravitational potential:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = [-\frac{\hbar^2}{2m}\nabla^2 + mV_G]\psi(x,t). \tag{2}$$

The gravitational interaction term $mV_G\psi(x,t)$ in the equation also suggests that the (passive gravitational) mass of the quantum system distributes throughout the whole region where its wave function $\psi(x,t)$ is nonzero, and the mass density of the system is $m|\psi(x,t)|^2$.

2.1.2 The answer of protective measurement

Protective measurement provides a more convincing argument for the existence of the mass and charge density distributions of a quantum system (Aharonov and Vaidman 1993; Aharonov, Anandan and Vaidman 1993; Aharonov, Anandan and Vaidman 1996). Like the conventional impulse measurement, it also uses the standard measuring procedure, but with an adiabatic coupling and an appropriate protection. This permits protective measurement to be able to measure the expectation values of observables on a single quantum system. In particular, the mass and charge density of a quantum system can be measured by protective measurement as expectation values of certain observables. For example, a protective measurement of the flux of the electric field of a charged quantum system out of a certain region will yield the expectation value of its charge inside this region, namely the integral of its charge density over this region. Similarly, we can also measure the mass density of a quantum system by a protective measurement of the flux of its gravitational field in principle (Anandan 1993).

As a typical example, consider a quantum system in a discrete nondegenerate energy eigenstate $\psi(x)$. In this case, the system itself supplies the protection of the state due to energy conservation and no artificial protection is needed. We take the measured observable A_n to be (normalized) projection operators on small spatial regions V_n having volume v_n:

$$A_n = \begin{cases} \frac{1}{v_n}, & \text{if } x \in V_n, \\ 0, & \text{if } x \notin V_n. \end{cases} \tag{3}$$

The protective measurement of A_n then yields

$$\langle A_n \rangle = \frac{1}{v_n}\int_{V_n} |\psi(x)|^2 dv = |\psi_n|^2, \tag{4}$$

where $|\psi_n|^2$ is the average of the density $\rho(x) = |\psi(x)|^2$ over the small region V_n. Then when $v_n \to 0$ and after performing measurements in sufficiently many regions V_n we can measure $\rho(x)$ everywhere in space.

Since the physical realization of the observable A_n must always resort to the electromagnetic or gravitational interaction between the measured system and the measuring device, what the above protective measurement measures is in fact the charge or mass density of the quantum system, and its result indicates that the mass and charge density is proportional to the modulus square of the wave function of the system, namely the density $\rho(x)$ (see Gao 2011 for a more detailed analysis).

In addition, it is worth noting that the expectation values of observables, which are measurable by protective measurements, are not the time-averaged properties of the evolution of a quantum system during a long period of time (Aharonov, Anandan and Vaidman 1996). Although a protective measurement cannot measure the expectation values at a precise instant, for an arbitrarily short period of time the measuring device always shifts by an amount proportional to the expectation value of the measured observable. Therefore, the mass and charge density of a single quantum system as expectation values of certain observables are the instantaneous properties of a quantum system, which are defined during an infinitesimal time interval at a given instant.

To sum up, protective measurement shows that a quantum system with mass m and charge Q, which is described by the wave function $\psi(x,t)$, has mass density $m|\psi(x,t)|^2$ and charge density $Q|\psi(x,t)|^2$, respectively.

2.2 The origin of the mass and charge density of a quantum system

We have argued that a charged quantum system has mass and charge density proportional to the modulus square of its wave function. In this section, we will further investigate the physical origin of the mass and charge density. Is it real or only effective? As we will see, the answer may provide an important clue to the physical meaning of the wave function.

2.2.1 The mass and charge density is effective

If the mass and charge density of a charged quantum system is real, that is, if the densities at different locations exist at the same time, then there will exist gravitational and electrostatic self-interactions of the density. Interestingly, the Schrödinger-Newton equation, which was proposed by Diósi (1984) and Penrose (1998), just describes the gravitational self-interaction of the mass

density. The equation for a single quantum system can be written as

$$i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t)-Gm^2\int\frac{|\psi(\mathbf{x}',t)|^2}{|\mathbf{x}-\mathbf{x}'|}d^3\mathbf{x}'\psi(\mathbf{x},t)+V\psi(\mathbf{x},t), \quad (5)$$

where m is the mass of the quantum system, V is an external potential, G is Newton's gravitational constant. As we will see below, although such gravitational self-interactions cannot yet be excluded by experiments, the existence of the electrostatic self-interaction for a charged quantum system already contradicts experimental observations.

If there is also an electrostatic self-interaction, then the equation for a free quantum system with mass m and charge Q will be

$$i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t)+(kQ^2-Gm^2)\int\frac{|\psi(\mathbf{x}',t)|^2}{|\mathbf{x}-\mathbf{x}'|}d^3\mathbf{x}'\psi(\mathbf{x},t). \quad (6)$$

Note that the gravitational self-interaction is attractive, while the electrostatic self-interaction is repulsive. It has been shown that the measure of the potential strength of the gravitational self-interaction is $\varepsilon^2=(\frac{4Gm^2}{\hbar c})^2$ for a free system with mass m (Salzman 2005). This quantity represents the strength of the influence of the self-interaction on the normal evolution of the wave function; when $\varepsilon^2\approx 1$ the influence is significant. Similarly, for a free charged system with charge Q, the measure of the potential strength of the electrostatic self-interaction is $\varepsilon^2=(\frac{4kQ^2}{\hbar c})^2$. As a typical example, for a free electron the potential strength of the electrostatic self-interaction will be $\varepsilon^2=(\frac{4ke^2}{\hbar c})^2\approx 1\times 10^{-3}$. This indicates that the electrostatic self-interaction will have a remarkable influence on the evolution of the wave function of a free electron[1]. If such an interaction indeed exists, it should have been detected by precise interference experiments on electrons. On the other hand, the superposition principle of quantum mechanics, which denies the existence of the observable electrostatic self-interaction, has been verified for microscopic particles with astonishing precision. As another example, consider the electron in a hydrogen atom. Since the potential of the electrostatic self-interaction is of the same order as the Coulomb potential produced by the nucleus, the energy levels of hydrogen atoms will be remarkably different from those predicted by quantum mechanics and confirmed by experiments.

Therefore, the mass and charge density of a quantum system cannot be real but be effective. This means that at every instant there is only a localized particle with the total mass and charge of the system, and during a time interval the time average of the ergodic motion of the particle forms the

[1]By contrast, the potential strength of the gravitational self-interaction for a free electron is $\varepsilon^2=(\frac{4Gm_e^2}{\hbar c})^2\approx 4\times 10^{-89}$.

effective mass and charge density[2]. There exist no gravitational and electrostatic self-interactions of the density in this case. Moreover, since protective measurement implies that the mass and charge density of a quantum system is an instantaneous property of the system, the ergodic motion of the particle must form the effective mass and charge density during an infinitesimal time interval (*not* during a finite time interval) at a given instant.

2.2.2 The ergodic motion of a particle is discontinuous

Which sort of ergodic motion then? If the ergodic motion of a particle is continuous, then it can only form the effective mass and charge density during a finite time interval, which contradicts the result of protective measurement. Thus it seems that the ergodic motion of a particle cannot be continuous. This is at least what the existing quantum mechanics says. However, there may exist a possible loophole here. Although the classical ergodic models that assume continuous motion are inconsistent with quantum mechanics due to the existence of a finite ergodic time, they may be not completely precluded by experiments if only the ergodic time is extremely short. After all quantum mechanics is also an approximation of a more fundamental theory of quantum gravity, in which there may exist a minimum time scale such as the Planck time. Therefore, we need to investigate the classical ergodic models more thoroughly.

Consider an electron in a one-dimensional box in the first excited state $\psi(x)$ (Aharonov and Vaidman 1993). Its wave function has a node at the center of the box, where its charge density is zero. Assume the electron performs a very fast continuous motion in the box, and during a very short time interval its motion generates an effective charge density distribution. Let's see whether this density can assume the same form as $e|\psi(x)|^2$, which is required by protective measurement[3]. Since the effective charge density is proportional to the amount of time the electron spends in a given position, the electron must be in the left half of the box half of the time and in the right half of the box half of the time. But it can spend no time at the center of the box where the effective charge density is zero; in other words, it must move at infinite velocity at the center. Certainly, the appearance of velocity faster than light or even infinite velocity may be not a fatal problem, as our discussion is entirely in the context of non-relativistic quantum mechanics,

[2]Even if there are only two masses and charges in space at a given instant, the densities formed by their motion also have gravitational and electrostatic interactions. Therefore, the mass and charge density of a quantum system can only be formed by the ergodic motion of one localized particle with the total mass and charge of the system.

[3]Note that in Nelson's stochastic mechanics the electron, which is assumed to undergo a Brownian motion, moves only within a region bounded by the nodes (Nelson 1966). This ensures that the theory can be equivalent to quantum mechanics in a limited sense. Obviously this sort of motion is not ergodic and cannot generate the required charge density distribution.

and especially the infinite potential in the example is also an ideal situation. However, it seems difficult to explain why the electron speeds up at the node and where the infinite energy required for the acceleration comes from. Moreover, the sudden acceleration of the electron near the node may also result in large radiation (Aharonov, Anandan and Vaidman 1993), which is inconsistent with the predictions of quantum mechanics. Again, it seems very difficult to explain why the accelerating electron does not radiate here.

Let's further consider an electron in a superposition of two energy eigenstates in two boxes $\psi_1(x) + \psi_2(x)$. In this example, even if one assumes that the electron can move with infinite velocity (e.g. at the nodes), it cannot *continuously* move from one box to another due to the restriction of the boxes. Therefore, any sort of continuous motion cannot generate the effective charge density $e|\psi_1(x) + \psi_2(x)|^2$. One may still object that this is merely an artifact of the idealization of infinite potential. However, even in this ideal situation, the model should also be able to generate the effective charge density by means of some sort of ergodic motion of the electron; otherwise it will be inconsistent with quantum mechanics. On the other hand, it is very common in quantum optics experiments that a single-photon wave packet is splitted into two branches moving along two well-separated paths in space. The wave function of the photon disappears outside the two paths for all practical purposes. Moreover, the experimental results are not influenced by the environment and setup between the two paths of the photon. Thus it is very difficult to imagine that the photon performs continuous ergodic motion back and forth in the space between its two paths.

In view of these serious drawbacks of the classical ergodic models and their inconsistency with quantum mechanics, we conclude that the ergodic motion of particles cannot be continuous but be discontinuous. If the motion of a particle is discontinuous, then the particle can readily move throughout all regions where the wave function is nonzero during an arbitrarily short time interval at a given instant. Furthermore, if the probability density of the particle appearing in each position is proportional to the modulus square of its wave function there at every instant, the discontinuous motion can also generate the right effective mass and charge density. This will solve the above problems plagued by the classical ergodic models. The discontinuous ergodic motion requires no existence of a finite ergodic time. Moreover, a particle undergoing discontinuous motion can also move from one region to another spatially separated region, no matter whether there is an infinite potential wall between them, and such discontinuous motion is not influenced by the environment and setup between these regions either. Besides, discontinuous motion can also solve the problems of infinite velocity and accelerating radiation. The reason is that no classical velocity and acceleration can be defined for discontinuous motion, and energy and momentum will require new definitions and understandings as in quantum mechanics.

In conclusion, we have argued that the mass and charge density of a quantum system, which can be measured by protective measurement, is not real but effective. Moreover, the effective mass and charge density is formed by the discontinuous motion of a localized particle, and the probability density of the particle appearing in each position is proportional to the modulus square of its wave function there. As a result, the wave function can be regarded as a description of the discontinuous motion of the localized particle. In the next subsection, we will give a detailed analysis of this suggested interpretation of the wave function.

2.2.3 An argument for randomness of discontinuous motion

Although the above analysis demonstrates that the ergodic motion of a particle is discontinuous, it doesn't say that the discontinuous motion must be random. In particular, the randomness of the result of a quantum measurement may be only apparent. In order to know whether the motion of particles is random or not, we need to analyze the cause of motion. For example, if motion has no *deterministic* cause, then it will be random, only determined by a probabilistic cause. This may also be the right way to find how particles move. Since motion involves change in position, if we can find the cause or instantaneous condition determining the change[4], we will be able to find how particles move in reality.

Let's consider the simplest states of motion of a free particle, for which the instantaneous condition determining the change of its position is a constant during the motion. In logic the instantaneous condition can only be deterministic or indeterministic. That the instantaneous condition is deterministic means that it leads to a deterministic change of the position of a particle at a given instant. That the instantaneous condition is indeterministic means that it only determines the probability of the particle appearing in each position in space at a given instant. If the instantaneous condition is deterministic, then the simplest states of motion of the free particle will have two possible forms. The first one is continuous motion with constant velocity, and the equation of motion of the particle is $x(t+dt) = x(t)+vdt$, where the deterministic instantaneous condition v is a contant[5]. The second one is discontinuous motion with infinite average velocity; the particle performs a finite jump along a fixed direction at every instant, where the jump distance is determined by the constant instantaneous condition[6]. On the other hand,

[4]The word "cause" used here only denotes a certain instantaneous condition determining the change of position, which may appear in the laws of motion. Our analysis is irrelevant to whether the condition has causal power or not.

[5]This deterministic instantaneous condition has been often called intrinsic velocity. It is different from the standard velocity, though they are equal in numerical values (Tooley 1988).

[6]In discrete space and time, the motion will be a discrete jump along a fixed direction at each time unit, and it will become continuous motion with constant velocity in the

if the instantaneous condition is indeterministic, then the simplest states of motion of the free particle will be random discontinuous motion with even position probability density. At each instant the probability density of the particle appearing in every position is the same.

Now let's see which sort of simplest states of motion are the solutions of the equation of free motion in quantum mechanics (i.e. the free Schrödinger equation). According to the analysis in the last subsection, the momentum eigenstates of a free particle, which are the solutions of the free Schrödinger equation, describe the discontinuous motion of the particle with even position probability density in space. Therefore, the simplest states of motion with a constant probabilistic instantaneous condition are the solutions of the equation of free motion, while the simplest states of motion with a constant deterministic instantaneous condition are not.

Then when assuming that (1) the simplest states of motion of a free particle are the solutions of the equation of free motion; and (2) the instantaneous condition determining the position change of a particle is always deterministic or indeterministic for any state of motion, the above result implies that motion, no matter it is free or forced, has no deterministic cause, and thus it is random and discontinuous, only determined by a probabilistic cause. This argument may be improved by further analyzing these two seemingly reasonable assumptions, but we will leave this for future work.

2.3 The wave function as a description of random discontinuous motion of particles

In classical mechanics, we have a clear physical picture of motion. It is well understood that the trajectory function $x(t)$ in classical mechanics describes the continuous motion of a particle. In quantum mechanics, the trajectory function $x(t)$ is replaced by a wave function $\psi(x,t)$. Since quantum mechanics is a more fundamental theory of the physical world, of which classical mechanics is only an approximation, it seems natural that the wave function should describe some sort of more fundamental motion of particles, of which continuous motion is only an approximation in the classical domain. The analysis in the last subsection provides a strong support for this conjecture, and it suggests that the wave function may indeed describe the more fundamental motion of particles, which is essentially discontinuous and random. In this section, we will give a more detailed analysis of this suggested interpretation of the wave function.

2.3.1 An analysis of random discontinuous motion of particles

The physical definition of random discontinuous motion of a particle is as follows. The position of the particle at each instant is only determined by

continuous limit.

a certain instantaneous condition at the instant in a probabilistic way, and this probabilistic instantaneous condition gives the probability density of the particle appearing in every position in space. As a result, the trajectory of the particle is essentially discontinuous, that is, its trajectory function, $x(t)$, is not continuous at every instant t[7]. Unlike the deterministic continuous motion, the trajectory function no longer provides a useful description for random discontinuous motion. In the following, we will give a strict description of random discontinuous motion based on the measure theory in mathematics. For simplicity we will mainly analyze the one-dimensional motion of particles, and the results can be readily extended to the three-dimensional situation.

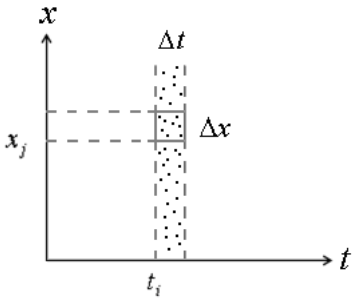

Fig.1 The description of random discontinuous motion of a single particle

We first analyze the random discontinuous motion of a single particle. Consider the state of motion of the particle in finite intervals Δt and Δx near a space-time point (t_i,x_j) as shown in Fig. 1. By the definition of random discontinuous motion, the positions of the particle form a random, discontinuous trajectory in this square region. We study the projection of this trajectory in the t-axis, which is a dense instant set in the time interval Δt. Let W be the discontinuous trajectory of the particle and Q be the square region $[x_j, x_j + \Delta x] \times [t_i, t_i + \Delta t]$. The dense instant set can be denoted by $\pi_t(W \cap Q) \in \Re$, where π_t is the projection on the t-axis. According to the measure theory, we can define the Lebesgue measure:

$$M_{\Delta x,\Delta t}(x_j, t_i) = \int_{\pi_t(W\cap Q)\in\Re} dt. \tag{7}$$

Since the sum of the measures of all such dense instant sets in the time interval Δt is equal to the length of the continuous time interval Δt, we have:

$$\sum_j M_{\Delta x,\Delta t}(x_j, t_i) = \Delta t. \tag{8}$$

[7]Recall that a function $x(t)$ is continuous if and only if for every t and every real number $\varepsilon > 0$, there exists a real number $\delta > 0$ such that whenever a point t_0 has distance less than δ to t, the point $x(t_0)$ has distance less than ε to $x(t)$.

Then we can define the measure density as follows[8]:

$$\rho(x,t) = \lim_{\Delta x, \Delta t \to 0} M_{\Delta x, \Delta t}(x,t)/(\Delta x \cdot \Delta t). \tag{9}$$

We call it position measure density or position density in brief. This quantity provides a strict description of the position distribution of the particle in an infinitesimal space interval dx near position x during an infinitesimal interval dt near instant t. In other words, $\rho(x,t)$ provides a strict description of the state of random discontinuous motion of the particle at instant t. From Eq. (8) we can see that $\rho(x,t)$ satisfies the normalization relation, namely $\int_{-\infty}^{+\infty} \rho(x,t)dx = 1$.

Since the position density will change with time in general, we can further define the position flux density $j(x,t)$ through the relation $j(x,t) = \rho(x,t)v(x,t)$, where $v(x,t)$ is the velocity of the local position density. It describes the change rate of the position density. Due to the conservation of measure, $\rho(x,t)$ and $j(x,t)$ satisfy the continuity equation:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial j(x,t)}{\partial x} = 0. \tag{10}$$

The position density $\rho(x,t)$ and position flux density $j(x,t)$ provide a complete description of the state of random discontinuous motion of a single particle.

It is direct to extend the description of the motion of a single particle to the motion of many particles. For the random discontinuous motion of N particles, we can define joint position density $\rho(x_1, x_2, ...x_N, t)$ and joint position flux density $j(x_1, x_2, ...x_N, t) = \rho(x_1, x_2, ...x_N, t)v(x_1, x_2, ...x_N, t)$. They also satisfy the continuity equation:

$$\frac{\partial \rho(x_1, x_2, ...x_N, t)}{\partial t} + \sum_{i=1}^{N} \frac{\partial j(x_1, x_2, ...x_N, t)}{\partial x_i} = 0. \tag{11}$$

When these N particles are independent, the joint position density can be reduced to the direct product of the position density for each particle, namely $\rho(x_1, x_2, ...x_N, t) = \prod_{i=1}^{N} \rho(x_i, t)$. Note that the joint position density $\rho(x_1, x_2, ...x_N, t)$ and joint position flux density $j(x_1, x_2, ...x_N, t)$ are not defined in the real three-dimensional space, but defined in the 3N-dimensional configuration space.

2.3.2 Interpreting the wave function

Although the motion of particles is essentially discontinuous and random, the discontinuity and randomness of motion is absorbed into the state of mo-

[8]The existence of this limit relies on the continuity of the evolution of the probabilistic instantaneous condition that determines the random discontinuous motion.

tion, which is defined during an infinitesimal time interval, by the descriptive quantities of position density $\rho(x,t)$ and position flux density $j(x,t)$. Therefore, the evolution of the state of random discontinuous motion of particles can be a deterministic continuous equation. By assuming that the non-relativistic equation of random discontinuous motion is the Schrödinger equation in quantum mechanics, both $\rho(x,t)$ and $j(x,t)$ can be expressed by the wave function in an unique way[9]:

$$\rho(x,t) = |\psi(x,t)|^2, \tag{12}$$

$$j(x,t) = \frac{\hbar}{2mi}[\psi^*(x,t)\frac{\partial\psi(x,t)}{\partial x} - \psi(x,t)\frac{\partial\psi^*(x,t)}{\partial x}]. \tag{13}$$

Correspondingly, the wave function $\psi(x,t)$ can be uniquely expressed by $\rho(x,t)$ and $j(x,t)$ (except for an absolute phase factor):

$$\psi(x,t) = \sqrt{\rho(x,t)}e^{im\int_{-\infty}^{x}\frac{j(x',t)}{\rho(x',t)}dx'/\hbar}. \tag{14}$$

In this way, the wave function also provides a complete description of the state of random discontinuous motion of particles[10]. For the motion of many particles, the joint position density and joint position flux density are defined in the 3N-dimensional configuration space, and thus the many-particle wave function, which is composed of these two quantities, is also defined in the 3N-dimensional configuration space.

Interestingly, we can reverse the above logic in some sense, namely by assuming the wave function is a complete description for the motion of particles, we can also reach the random discontinuous motion of particles, independent of our previous analysis. If the wave function $\psi(x,t)$ is a description of the state of motion for a single particle, then the quantity $|\psi(x,t)|^2dx$ not only gives the probability of the particle being found in an infinitesimal space interval dx near position x at instant t (as in standard quantum mechanics), but also gives the objective probability of the particle being there. This accords with the common-sense assumption that the probability distribution of the measurement results of a property is the same as the objective distribution of the property in the measured state. Then at instant t the particle may appear in any location where the probability density $|\psi(x,t)|^2$ is nonzero, and during an infinitesimal time interval near instant t the particle will move throughout the whole region where the wave function $\psi(x,t)$ spreads. Moreover, its position density will be equal to the probability

[9]Note that the relation between $j(x,t)$ and $\psi(x,t)$ depends on the concrete evolution under an external potential such as electromagnetic vector potential. By contrast, the relation $\rho(x,t) = |\psi(x,t)|^2$ holds true universally, independent of the concrete evolution.

[10]The picture of random discontinuous motion may exist not only for position but also for other dynamical variables such as momentum and energy, and thus this interpretation may also apply to the wave functions in momentum space etc.

density $|\psi(x,t)|^2$. Obviously this kind of motion is essentially random and discontinuous.

One important point needs to be pointed out here. Since the wave function in quantum mechanics is defined at an instant, not during an infinitesimal time interval, it should be regarded not simply as a description of the state of random discontinuous motion of particles, but as a description of the instantaneous condition or instantaneous intrinsic property of the particles that determines their random discontinuous motion, at a deeper level [11]. In particular, the modulus square of the wave function *determines* the probability density of the particles appearing in every position in space at a given instant. This intrinsic property may be called indeterministic disposition or propensity[12]. By contrast, the position density $\rho(x,t)$ and position flux density $j(x,t)$, which are defined during an infinitesimal time interval, are only a description of the state of the resulting random discontinuous motion of particles, and they are determined by the wave function. In this sense, we may say that the motion of particles is "guided" by their wave function in a probabilistic way.

The suggested interpretation of the wave function in terms of random discontinuous motion of particles might be taken as a natural realistic extension of the orthodox view. The naturalness of the extension lies in that it still makes particles ontological and the wave function epistemological[13]. That the extension is realistic is obvious. According to Born's probability interpretation, the modulus square of the wave function of a particle gives the probability density of the particle *being found* in certain positions, while according to the suggested interpretation, the modulus square of the wave function also gives the objective probability density of the particle *being* there. Certainly, the transition process from "being" to "being found", which is closely related to the measurement problem, needs to be further explained. We will discuss this important issue in Section 4.

3 Schrödinger's equation and the conservation laws

After investigating the physical meaning of the wave function, we will further analyze its linear evolution. Many quantum mechanics textbooks provide a heuristic "derivation" of the Schrödinger equation of the wave function.

[11] From a logical point of view, for the random discontinuous motion of particles, the particles should also have an intrinsic property that determines their discontinuous motion in a probabilistic way, otherwise they would not "know" how frequently they should appear in every position in space. See also the definition of random discontinuous motion given in the last subsection.

[12] It is worth noting that this kind of propensity relates to the objective motion of particles, not to the measurement on the particles (cf. Suárez 2004).

[13] By contrast, the de Broglie-Bohm theory and the many-worlds interpretation both attach reality to the wave function itself (Bohm 1952; Everett 1957).

It begins with the assumption that the state of a free quantum system has the form of a plane wave $e^{i(kx-\omega t)}$. When combining with the de Broglie relations for momentum and energy $p = \hbar k$ and $E = \hbar\omega$, this state becomes $e^{i(px-Et)/\hbar}$. Then it uses the non-relativistic energy-momentum relation $E = p^2/2m$ to obtain the free particle Schrödinger equation. Lastly, this equation is generalized to include an external potential, and the end result is the Schrödinger equation.

In this section, we will show that the heuristic "derivation" of the free Schrödinger equation can be turned into a real derivation by resorting to spacetime translation invariance and relativistic invariance. Spacetime translation gives the definitions of momentum and energy, and spacetime translation invariance entails that the state of a free quantum system or an isolated system with definite momentum and energy assumes the plane wave form $e^{i(px-Et)/\hbar}$. Besides, the relativistic invariance of the free states further determines the relativistic energy-momentum relation, which non-relativistic approximation is $E = p^2/2m$. The new integrated analysis may help to undertsand the physical origin of the Schrödinger equation, and moreover, it is also helpful for understanding momentum and energy and their conservation for random discontinuous motion of particles.

3.1 Spacetime translation and its invariance

We will first show that the free states of motion for a quantum system can be basically determined by spacetime translation invariance. The spacetime translation invariance of natural laws reflects the homogeneity of space and time. The homogeneity of space ensures that the same experiment performed at two different places gives the same result, and the homogeneity in time ensures that the same experiment repeated at two different times gives the same result. There are in general two different pictures of translation: active transformation and passive transformation. The active transformation corresponds to displacing the studied system, and the passive transformation corresponds to moving the environment (the coordinate system etc). Physically, the equivalence of the active and passive pictures is due to the fact that moving the system one way is equivalent to moving the environment the other way by an equal amount In the following, we will mainly analyze spacetime translations in terms of active transformations.

A space translation operator can be defined as

$$T(a)\psi(x,t) = \psi(x-a,t). \tag{15}$$

It means translating (without distortion) the state of a system, $\psi(x,t)$, by an amount a in the positive x direction. The operator preserves the norm of the state because $\int_{-\infty}^{\infty}\psi^*(x,t)\psi(x,t)dx = \int_{-\infty}^{\infty}\psi^*(x-a,t)\psi(x-a,t)dx$. This implies that $T(a)$ is unitary, satisfying $T^\dagger(a)T(a) = I$. As a unitary

operator, $T(a)$ can be further expressed as

$$T(a) = e^{-iaP}, \tag{16}$$

where P is called the generator of space translation, and it is Hermitian and its eigenvalues are real. By expanding $\psi(x-a,t)$ in order of a, we can further get

$$P = -i\frac{\partial}{\partial x}. \tag{17}$$

Similarly, a time translation operator can be defined as

$$U(t)\psi(x,0) = \psi(x,t). \tag{18}$$

Let the evolution equation of state be the following form:

$$i\frac{\partial\psi(x,t)}{\partial t} = H\psi(x,t). \tag{19}$$

where H is a to-be-determined operator that depends on the properties of the system. In the following analysis of this subsection, we assume H is independent of the evolved state, namely the evolution is linear[14]. Then the time translation operator $U(t)$ can be expressed as $U(t) = e^{-itH}$, and H is the generator of time translation. Note that we cannot determine whether $U(t)$ is unitary and H is Hermitian here.

Let's now analyze the implications of spacetime translation invariance for the law of motion of an isolated system. First, time translational invariance requires that H has no time dependence, namely $dH/dt = 0$. It is worth stressing that the linearity of evolution is an important presupposition of this result. If H depends on the state, then obviously we cannot obtain $dH/dt = 0$ because the state is related to time, though the state-dependent H also satisfies the time translational invariance. Secondly, space translational invariance requires $[T(a), U(t)] = 0$, which further leads to $[P, H] = 0$[15]. Again, we stress that the linearity of evolution is an important presupposition of this result. If $U(t)$ depends on the state, then the space translational invariance will only lead to $U(t, T\psi)T(a)\psi(x,0) = T(a)U(t,\psi)\psi(x,0)$, from which we cannot obtain $[T(a), U(t)] = 0$.

When $dH/dt = 0$, the solutions of the evolution equation Eq.(19) assume the following form

$$\psi(x,t) = \varphi_E(x)e^{-iEt}, \tag{20}$$

[14]This is an important presupposition in our derivation. We will consider the possible case of nonlinearity of H in the next subsection.

[15]See Shankar (1994) for a clear derivation of these two formulae.

where E is a constant, and $\varphi_E(x)$ is the eigenstate of H and satisfies the time-independent equation:

$$H\varphi_E(x) = E\varphi_E(x). \tag{21}$$

The commutative relation $[P, H] = 0$ further implies that P and H have common eigenstates. This means that $\varphi_E(x)$ is also the eigenstate of P. Since the eigenstate of $P = -i\frac{\partial}{\partial x}$ is e^{ipx}, where p is a real eigenvalue, the solution of the evolution equation Eq.(19) for an isolated system will be $e^{i(px-Et)}$. In quantum mechanics, P and H, the generators of space translation and time translation, are also called momentum operator and energy opertaor, respectively. Correspondingly, $e^{i(px-Et)}$ is the eigenstate of both momentum and energy, and p and E are the corresponding momentum and energy eigenvalues, respectively. In other words, the state $e^{i(px-Et)}$ describes an isolated system (e.g. a free microscopic particle) with definite momentum p and energy E.

3.2 Relativistic invariance

The relation between momentum p and energy E can be determined by the relativistic invariance of the momentum eigenstate $e^{i(px-Et)}$, and it turns out to be $E^2 = p^2c^2 + m^2c^4$, where m is the mass of the system, and c is the speed of light. In the non-relativistic domain, the energy-momentum relation reduces to $E = p^2/2m$.

Now we will derive the relation between momentum p and energy E in the relativistic domain. Consider two inertial frames S_0 and S with coordinates x_0, t_0 and x, t. S_0 is moving with velocity v relative to S. Then x, t and x_0, t_0 satisfy the Lorentz transformations:

$$x_0 = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \tag{22}$$

$$t_0 = \frac{t - xv/c^2}{\sqrt{1 - v^2/c^2}}. \tag{23}$$

Suppose the state of a free particle is $\psi = e^{i(p_0x_0 - E_0t_0)}$, an eigenstate of P, in S_0, where p_0, E_0 is the momentum and energy of the particle in S_0, respectively. When described in S by coordinates x, t, the state is

$$\psi = e^{i(p_0\frac{x-vt}{\sqrt{1-v^2/c^2}} - E_0\frac{t-xv/c^2}{\sqrt{1-v^2/c^2}})} = e^{i(\frac{p_0+E_0v/c^2}{\sqrt{1-v^2/c^2}}x - \frac{E_0+p_0v}{\sqrt{1-v^2/c^2}}t)}. \tag{24}$$

This means that in frame S the state is still the eigenstate of P, and the

corresponding momentum p and energy E is[16]

$$p = \frac{p_0 + E_0 v/c^2}{\sqrt{1 - v^2/c^2}}, \tag{25}$$

$$E = \frac{E_0 + p_0 v}{\sqrt{1 - v^2/c^2}}. \tag{26}$$

We further suppose that the particle is at rest in frame S_0. Then the velocity of the particle is v in frame S[17]. Considering that the velocity of a particle in the momentum eigenstate $e^{i(px-Et)}$ or a wavepacket superposed by these eigenstates is defined as the group velocity of the wavepacket, namely

$$u = \frac{dE}{dp}, \tag{27}$$

we have

$$dE_0/dp_0 = 0, \tag{28}$$

$$dE/dp = v. \tag{29}$$

Eq.(28) means that E_0 and p_0 are independent. Moreover, since the particle is at rest in S_0, E_0 and p_0 do not depend on v. By differentiating both sides of Eq.(25) and Eq.(26) relative to v, we obtain

$$\frac{dp}{dv} = \frac{v}{c^2}\frac{p_0 + E_0 v/c^2}{(1 - v^2/c^2)^{\frac{3}{2}}} + \frac{E_0/c^2}{(1 - v^2/c^2)^{\frac{1}{2}}}, \tag{30}$$

$$\frac{dE}{dv} = \frac{v}{c^2}\frac{E_0 + p_0 v}{(1 - v^2/c^2)^{\frac{3}{2}}} + \frac{p_0}{(1 - v^2/c^2)^{\frac{1}{2}}}. \tag{31}$$

Dividing Eq.(31) by Eq.(30) and using Eq.(29) we obtain

$$\frac{p_0}{\sqrt{1 - v^2/c^2}} = 0. \tag{32}$$

This means that $p_0 = 0$. Inputing this important result into Eq.(26) and Eq.(25), we immediately obtain

$$E = \frac{E_0}{\sqrt{1 - v^2/c^2}}, \tag{33}$$

[16] Alternatively we can obtain the transformations of momentum and energy by directly requiring the relativistic invariance of momentum eigenstate $e^{i(px-Et)}$, which leads to the relation $px - Et = p_0 x_0 - E_0 t_0$.

[17] We can also get this result from the definition Eq. (27) by using the above transformations of momentum and energy Eq.(25) and Eq.(26).

$$p = \frac{E_0 v/c^2}{\sqrt{1 - v^2/c^2}}. \tag{34}$$

Then the energy-momentum relation is[18]:

$$E^2 = p^2c^2 + E_0^2, \tag{35}$$

where E_0 is the energy of the particle at rest, called rest energy of the particle, and p and E is the momemtum and energy of the particle with velocity v. By defining $m = E_0/c^2$ as the mass of the particle[19], we can further obtain the familar energy-momentum relation

$$E^2 = p^2c^2 + m^2c^4. \tag{36}$$

In the non-relativistic domain, this energy-momentum relation reduces to $E = p^2/2m$.

3.3 Derivation of the free Schrödinger equation

The relation between energy E and momentum p for momentum eigenstates in the non-relativistic domain implies that the operator relation is $H = P^2/2m$ for an isolated system, where H is the free Hamiltonian of the system. Note that since the value of E is real by Eq.(35), H is Hermitian and $U(t)$ is unitary for free evolutions. By inputing this operator relation into the evolution equation Eq.(19), we can obtain the free evolution equation, which assumes the same form as the free Schrödinger equation in quantum mechanics[20]:

$$i\frac{\partial\psi(x,t)}{\partial t} = -\frac{1}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}. \tag{37}$$

It is worth noting that, unlike the free Schrödinger equation, the reduced Planck constant $\hbar$ with dimension of action is missing in this equation. However, this is in fact not a problem. The reason is that the dimension of $\hbar$ can be absorbed in the dimension of the mass m in principle. For example, we can stipulate the dimensional relations as $p = 1/L$, $E = 1/T$ and $m = T/L^2$, where L and T represents the dimensions of space and time, respectively (see Duff, Okun and Veneziano 2002 for more discussions).

[18]Most existing "derivations" of the energy-momentum relation are based on the somewhat complex analysis of an elastic collision process. Moreover, they resort to either the Newtonian limits (e.g. $p = mv$) or less fundamental relations (e.g. $p = Eu/c^2$) or even some mathematical intuition (e.g. four-vectors) (see Sonego and Pin 2005 and references therein).

[19]According to the analysis here, it seems that we can in principle avoid talking about mass in modern physics from a more fundamental view (cf. Okun 2009).

[20]This also means that the Klein-Gordon equation can be derived in the relativistic domain.

Moreover, the value of $\hbar$ can be set to the unit of number 1 in principle. Thus the above equation is essentially the free Schrödinger equation in quantum mechanics.

By using the definition of classical potential and requiring an appropriate expectation value correspondence, $d < P > /dt =< F >=< \partial V/\partial x >$, we can further obtain (*not* derive) the Schrödinger equation under an external potential $V(x,t)$[21]:

$$i\frac{\partial\psi(x,t)}{\partial t} = -\frac{1}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x,t)\psi(x,t). \tag{38}$$

The general form of a classical potential may be $V(x, \frac{\partial}{\partial x}, t)$, and its concrete form is determined by the non-relativistic approximation of the quantum interactions involved, which are described by the relativistic quantum field theory. Since the potential $V(x,t)$ is real-valued, the Hamiltonian $H = P^2/2m + V(x,t)$ is Hermitian, and as a result, the time translation operator or evolution operator $U(t)$ is also unitary.

3.4 Further discussions

We have derived the free Schrödinger equation in quantum mechanics based on spacetime translation invariance and relativistic invariance. The derivation may not only make the equation more logical and understandable, but also shed some new light on the physical meaning of the wave function $\psi(x,t)$ in the equation.

The free Schrödinger equation is usually "derived" in textbooks by analogy and correspondence with classical physics. There are at least two mysteries in such a heuristic "derivation". First, even if the behavior of microscopic particles likes wave and thus a wave function is needed to describe them, it is unclear why the wave function must assume a complex form. Indeed, when Schrödinger originally invented his equation, he was very puzzled by the inevitable appearance of the imaginary unit "i" in the equation. Next, one doesn't know why there are the de Broglie relations for momentum and energy and why the non-relativistic energy-momentum relation must be $E = p^2/2m$. Usually one can only resort to experience and classical physics to answer these questions. This is unsatisfactory in logic as quantum mechanics is a more fundamental theory, of which classical mechanics is only an approximation.

As we have argued above, the key to unveil these mysteries is to analyze the origin of momentum and energy. According to the modern understanding, spacetime translation gives the definitions of momentum and energy. The momentum operator P is defined as the generator of space translation, and it is Hermitian and its eigenvalues are real. Moreover, the form

[21] In order to derive the complete Schrödinger equation, we need a fundamental theory of interactions.

of momentum operator can be uniquely determined by its definition. It is $P = -i\frac{\partial}{\partial x}$, and its eigenstate is e^{ipx}, where p is a real eigenvalue. Similarly, the energy operator H is defined as the generator of time translation. But its form is determined by the concrete situation. Fortunately, for an isolated system the form of energy operator, which determines the evolution equation, can be fixed by the requirements of spacetime translation invariance and relativitic invariance (when assuming the evolution is linear). Concretely speaking, time translational invariance requires that $dH/dt = 0$, and the solution of the evolution equation $i\frac{\partial\psi(x,t)}{\partial t} = H\psi(x,t)$ must assume the form $\psi(x,t) = \varphi_E(x)e^{-iEt}$. Besides, space translational invariance requires $[P, H] = 0$, and this further determines that $\varphi_E(x)$ is the eigenstate of P, namely $\varphi_E(x) = e^{ipx}$. Thus spacetime translation invariance entails that the state of an isolated system with definite momentum and energy assumes the plane wave form $e^{i(px-Et)}$. Furthermore, the relation between p and E or the energy-momentum relation can be determined by the relativistic invariance of the momentum eigenstate $e^{i(px-Et)}$, and its non-relativistic approximation is just $E = p^2/2m$. Then we can obtain the form of energy operator for an isolated system, $H = P^2/2m$, and the free Schrödinger equation, Eq.(37). To sum up, this analysis may answer why the wave function must assume a complex form in general and why there are the de Broglie relations and why the non-relativistic energy-momentum relation is what it is.

So far so good. But how does the wave function $\psi(x,t)$ in the thus-derived free Schrödinger equation relate to the actual physical state of the system? Without answering this question the above analysis seems vacuous in physics. This leads us to the problem of interpreting the wave function. According to the standard probability interpretation, the wave function in quantum mechanics is a probability amplitude, and its modulus square gives the probability density of finding a particle in certain locations. Notwithstanding the success of the standard interpretation, our derivation of the free Schrödinger equation seems to suggest that the wave function $\psi(x,t)$ is a description of the objective physical state of a quantum system, rather than the probability amplitude relating only to measurement. In our derivation we never refer to the measurement of the isolated system after all. Moreover, the derivation seems to further suggest that the wave function $\psi(x,t)$ is a complete description of the physical state of the system. As we have argued in the last subsection, $\psi(x,t)$ can be regarded as a complete description of the state of random discontinuous motion of a particle, and $|\psi(x,t)|^2dx$ gives the objective probability of the particle being in an infinitesimal space interval dx near position x at instant t. This objective interpretation of the wave function is quite consistent with the above derivation of the free Schrödinger equation.

On the other hand, the derivation may provide a further argument for the non-existence of continuous motion from the aspect of the laws of motion.

Continuous motion can be regarded as a very special form of discontinuous motion, for which the position density of a particle is $\rho(x,t) = \delta^2(x - x(t))$ and its velocity is $v(t) = dx(t)/dt$, where $x(t)$ is the continuous trajectory of the particle. However, such states are not solutions of the free Schrödinger equation, though they do satisfy the continuity equation. According to the free Schrödinger equation, an initial local state like $\delta(x - x_0)$ cannot sustain its locality during the evolution, and it will immediately spread throughout the whole space. Thus the law of free motion, which is derived based on the requirements of spacetime translation invariance etc, seems to imply that the motion of a particle cannot be continuous but be essentially discontinuous. Note that our derivation of the free Schrödinger equation does not depend on the picture of discontinuous motion, and thus this argument for the non-existence of continuous motion is not a vicious circle.

Lastly, it is worth stressing again that the linearity of evolution (i.e. that the Hamiltonian H is independent of the evolved state) is an important presupposition in our derivation of the free Schrödinger equation. It can be reasonably assumed that the linear evolution and nonlinear evolution both exist, and moreover, they satisfy spacetime translation invariance respectively because they cannot counteract each other in general. Then our derivation only shows that the linear part of free evolution, if satisfying spacetime translation invariance and relativistic invariance, must assume the same form as the free Schrödinger equation in the non-relativistic domain. Obviously, our derivation cannot exclude the existence of nonlinear quantum evolution. Moreover, since a general nonlinear evolution can readily satisfy the spacetime translation invariance, the invariance requirement can no longer determine the concrete form of possible nonlinear evolution.

3.5 On the conservation of energy and momentum

The conservation of energy and momentum is one of the most important principles in modern physics. In this subsection, we will analyze the physical basis and meaning of this principle in quantum mechanics, especially its relationship with the linearity of quantum dynamics.

As we have noted in the above derivation of the free Schrödinger equation, the origin of momentum and energy is closely related to spacetime translation; the momentum operator P and energy operator H are defined as the generators of space translation and time translation, respectively. Moreover, it is well known that the conservation of energy and momentum results from spacetime translation invariance. The usual derivation is as follows. The evolution law for an isolated system satisfies spacetime translation invariance due to the homogeneity of space and time. Time translational invariance requires that H has no time dependence, namely $dH/dt = 0$, and space translational invariance requires that the generators of space translation and time translation are commutative, namely $[P, H] = 0$. Then by

Ehrenfest's theorem for an arbitrary observable A

$$\frac{d\langle A\rangle}{dt} = \langle\frac{\partial A}{\partial t}\rangle - i\langle[A, H]\rangle, \tag{39}$$

where $\langle A\rangle = \int \psi^*(x,t)A\psi(x,t)dx$ is defined as the expectation value of A, we have

$$\frac{d\langle H\rangle}{dt} = 0, \tag{40}$$

and

$$\frac{d\langle P\rangle}{dt} = 0. \tag{41}$$

This means that the expectation values of energy and momentum are conserved. Moreover, for arbitrary functions $f(H)$ and $f(P)$, we also have

$$\frac{d\langle f(H)\rangle}{dt} = 0, \tag{42}$$

and

$$\frac{d\langle f(P)\rangle}{dt} = 0. \tag{43}$$

This is equivalent to the constancy of the expectation values of the generating functions or spacetime translation operators $U(a) \equiv e^{-iaH}$ and $T(a) \equiv e^{-iaP}$

$$\frac{d\langle U(a)\rangle}{dt} = 0, \tag{44}$$

and

$$\frac{d\langle T(a)\rangle}{dt} = 0. \tag{45}$$

By these two equations it follows that the probability distributions of energy eigenvalues and momentum eigenvalues are constant in time. This statement is usually defined as the conservation of energy and momentum in quantum mechanics.

Now let's analyze the implications of this derivation for the meaning of the conservation of energy and momentum. First of all, we point out that the linearity of evolution is an indispensable presupposition in the derivation. As we have stressed in the derivation of the free Schrödinger equation, spacetime translation invariance does not lead to $dH/dt = 0$ and $[P, H] = 0$ without assuming the linearity of evolution. Therefore, the common wisdom that invariance or symmetry implies laws of conservation only holds true for linear evolutions; spacetime translation invariance no longer leads

to the conservation of energy and momentum for any nonlinear evolution, and the invariance imposes no restriction for the nonlinear evolution either. Moreover, for a general nonlinear evolution $H(\psi)$, energy and momentum will be not conserved by Ehrenfest's theorem[22]:

$$\frac{d\langle H(\psi)\rangle}{dt} = \langle\frac{\partial H(\psi)}{\partial t}\rangle - i\langle[H(\psi), H(\psi)]\rangle = \langle\frac{\partial H(\psi)}{\partial t}\rangle \neq 0, \tag{46}$$

$$\frac{d\langle P\rangle}{dt} = \langle\frac{\partial P}{\partial t}\rangle - i\langle[P, H(\psi)]\rangle = -i\langle[P, H(\psi)]\rangle = -\langle\frac{\partial H(\psi)}{\partial x}\rangle \neq 0. \tag{47}$$

We can see the violation of the conservation of energy and momentum more clearly by analyzing the nonlinear evolution of momentum eigenstates $e^{i(px-Et)}$ and their superpositions. If a nonlinear evolution can conserve energy and momentum for momentum eigenstates, then the momentum eigenstates must be the solutions of the nonlinear evolution equation; otherwise the evolution will change the definite momentum eigenvalues or energy eigenvalues or both, and thus the conservation of energy and momentum will be violated. Some nonlinear evolutions can satisfy this requirement. For example, when $H(\psi) = P^2/2m + \alpha|\psi|^2$, the solutions still include the momentum eigenstates $e^{i(px-Et)}$, where $E = p^2/2m + \alpha$, and thus energy and momentum are conserved for such nonlinear evolutions of momentum eigenstates. However, even if a nonlinear evolution can conserve energy and momentum for momentum eigenstates, it cannot conserve energy and momentum for the superpositions of momentum eigenstates. The reason is obvious. Only for a linear evolution the momentum eigenstates and their superpositions can both be the solutions of the evolution equation. For any nonlinear evolution $H(\psi)$, if the momentum eigenstates are already its solutions, then their linear superpositions cannot be its solutions. This means that the coefficients of the momentum eigenstates in the superposition will change with time during the evolution. The change of amplitudes of the coefficients directly leads to the change of the probability distribution of momentum eigenvalues and energy eigenvalues, while the change of phases of the coefficients leads to the change of the momentum eigenvalues or energy eigenvalues, which also leads to the change of the probability distribution of momentum eigenvalues or energy eigenvalues. In fact, a nonlinear evolution may not only change the probability distributions of energy and momentum eigenvalues, but also change the energy-momentum relation in general cases (e.g. in the above example)[23]. These results are understandable when considering the fact that

[22]In order to ensure that the nonlinear evolution is unitary and thus the total probability is conserved in time, the Hamiltonian $H(\psi)$ must be also Hermitian. Besides, this property is also required to ensure that the energy eigenvalues (which satisfy the equation $H(\psi)\psi(x) = E\psi(x)$) are real. When the Hamiltonian $H(\psi)$ is Hermitian, the Ehrenfest theorem still holds true.

[23]This will violate the relativistic invariance of momentum eigenstates.

a nonlinear evolution of the spatial wave function will generally introduce a time-dependent interaction between its different momentum eigenstates, which is equivalent to introducing a time-dependent external potential for its free evolution in some sense. Therefore, it is not beyond expectation that a nonlinear evolution violates the conservation of energy and momentum in general.

Two points needs to be stressed here. First, energy and momentum are still defined as usual for nonlinear evolutions in the above discussions. One may object that they should be re-defined for a nonlinear evolution. However, this may be not the case. The reason is as follows. Momentum is defined as the generator of space translation, and this definition uniquely determines that its eigenstates are e^{ipx}. Similarly, energy is defined as the generator of time translation, and this definition uniquely determines that its eigenstates satisfy $H(\psi)\psi(x) = E\psi(x)$. Since these definitions are independent of whether the evolution of the state is linear or nonlinear, they should have a fundamental status in any theory formulated in space and time such as quantum mechanics. The second point is that the above arguments implicitly assumes that the nonlinear evolution $H(\psi)$ is universal, i.e., that it applies to all possible states. If the nonlinear evolution only applies to some special states, then the evolution may still conserve energy and momentum. For example, suppose the nonlinear evolution $H(\psi) = P^2/2m + \alpha|\psi|^2$ applies only to the momentum eigenstates $e^{i(px-Et)}$ and the linear evolution $H(\psi) = P^2/2m$ applies to the superpositions of momentum eigenstates, then energy and momentum are still conserved during the evolution. On the other hand, it has been argued that the universal nonlinear quantum dynamics has a serious drawback, namely that the description of composite systems depends on a particular basis in a Hilbert space (Czachor 1995). If a nonlinear quantum evolution only applies to certain privileged bases due to some reason, then such nonlinear quantum dynamics may be logically consistent and also conserve energy and momentum (Gao 2004).

The second implication of the above derivation of the conservation laws is that spacetime translation invariance implies the conservation of energy and momentum for individual states, *not* for an ensemble of identical systems. As in the derivation of the free Schrödinger equation, we only refer to an isolated system and never refer to any ensemble of identical systems in the derivation of the conservation laws. Moreover, the transformations of spacetime translation also apply to a single isolated system. Therefore, what the derivation tells us is that spacetime translation invariance implies the conservation of energy and momentum for the linear evolution of the states of an isolated system. The conservation of energy and momentum for a single system means that the objective probability distributions of energy eigenvalues and momentum eigenvalues are constant during the evolution of the state of the system. As argued before, the objective probability can be well understood according to the suggested interpretation of the wave func-

tion in terms of random discontinuous motion. Similarly, our analysis of nonlinear evolutions also shows that a universal nonlinear evolution violates the conservation of energy and momentum for individual systems.

This implication raises a further issue. It is well known that the conservation of energy and momentum in quantum mechanics refers to an ensemble of identical systems, not to individual systems, and its precise statement is that the probability distributions of the measurement results of energy and momentum for an ensemble of identical isolated systems are the same at every instant during the evolution of the systems in the ensemble. But as we have argued above, the derivation of the conservation laws based on spacetime translation invariance is for individual isolated systems, not for an ensemble of these systems. The derivation never refers to the measurements of these system either. Therefore, there is still a gap (which maybe very large) between the derivation and the conservation laws in quantum mechanics. Undoubtedly we must analyze the measurement process in order to fill the gap. We will postpone the detailed analysis of the measurement problem to the next section. Here we only want to answer a more general question. If the conservation laws in quantum mechanics are indeed valid as widely thought, then what are their implications for the evolution of individual states?

First of all, the evolution of the state of an isolated system cannot contain a universal deterministic nonlinear evolution, which applies to all possible states; otherwise the evolution will violate the conservation of energy and momentum not only at the individual level but also at the ensemble level. Next, the evolution may contain linear evolutions as well as special deterministic nonlinear evolutions that apply only to certain privileged states. They can both conserve energy and momentum for individual states[24]. Lastly, the evolution may also contain a (universal) stochastic nonlinear evolution, which applies to all possible states. Although the evolution cannot conserve energy and momentum for individual states, it may conserve energy and momentum for an ensemble of identical states. As we will see in the next section, the collapse of the wave function may be such a stochastic nonlinear evolution.

To summarize, we have analyzed the relationships between the conservation of energy and momentum, spacetime translation invariance and the linearity of quantum dynamics. It has been often claimed that the conservation of energy and momentum is a geometry conservation law, resulting from the requirement of spacetime translation invariance. However, this common-sense view is not wholly right. Only when assuming the linearity of quantum dynamics, can spacetime translation invariance lead to the conservation of energy and momentum. Moreover, the connection between

[24]For more discussions about the arguments for linear quantum dynamics see Holman (2006) and references therein.

invariance of natural laws and conservation laws is for individual states, not for an ensemble of identical states. Although a nonlinear evolution of the wave function can readily satisfy spacetime translation invariance, the invariance can no longer lead to the conservation of energy and momentum, let alone determining the form of the nonlinear evolution. Rather, a universal nonlinear evolution that applies to all possible states will inevitably violate the conservation of energy and momentum.

Since the conservation of energy and momentum is required by spacetime translation invariance only for the linear evolution of the wave function of an isolated system, the principle cannot exclude the existence of a possible nonlinear evolution that may violate it. In other words, spacetime translation invariance is no longer the reason to require that the evolution of the wave function of an isolated system must conserve energy and momentum. On the other hand, the conservation of energy and momentum may still hold true for an ensemble of identical isolated systems as claimed by the standard quantum mechanics. Therefore, a (universal) stochastic nonlinear evolution of the wave function may exist. Although such evolutions cannot conserve energy and momentum for individual states, it may conserve energy and momentum at the ensemble level. However, unlike the linear evolution, which is natural in the sense that its form can be uniquely determined by the invariance requirements, the stochastic nonlinear evolution must have a physical origin, and its form can only be detertemined by the underlying mechanism. In the next section, we will investigate the possible stochastic nonlinear evolution of the wave function.

4 A suggested solution to the measurement problem

In standard quantum mechanics, it is postulated that when a wave function is measured by a macroscopic device, it will no longer follow the linear Schrödinger equation, but instantaneously collapse to one of the wave functions that correspond to definite measurement results. However, this collapse postulate is only a makeshift, and the theory does not tell us why and how the definite measurement result appears (Bell 1990). There are in general two ways to solve the measurement problem. The first one is to integrate the collapse evolution with the normal Schrödinger evolution into a unified dynamics, e.g. in the dynamical collapse theories (Ghirardi 2008). The second way is to reject the postulate and assume that the Schrödinger equation completely describes the evolution of the wave function. There are two main alternative theories along this avoiding-collapse direction. The first one is the de Broglie-Bohm theory (de Broglie 1928; Bohm 1952), which takes the wave function as an incomplete description and adds some hidden variables to explain the emergence of definite measurement results. The second one

is the many-worlds interpretation (Everett 1957), which assumes the existence of many worlds to explain our definite experience in this world and still regards the wave function as a complete description of the whole worlds. In this section, we will analyze the possible implications of our suggested interpretation of the wave function for these solutions to the measurement problem. It turns out that the suggested interpretation disfavors the de Broglie-Bohm theory and the many-worlds interpretation but favors the dynamical collapse theories, which regard the wavefunction collapse as a real and dynamical physical process.

At first sight, the above three theories seem apparently inconsistent with the suggested interpretation of the wave function. They all attach reality to the wave function, e.g. taking the wave function as a real physical entity on configuration space or assuming the wave function has a field-like spatiotemporal manifestation in the real three-dimensional space (see, e.g. Ghirardi 1997, 2008; Wallace and Timpson 2009). But according to our suggested interpretation, the wave function is not a field-like physical entity on configuration space[25]; rather, it is a description of the random discontinuous motion of particles in real space (and at a deeper level a description of the instantaneous intrinsic property of the particles that determines their random discontinuous motion). Anyway, in spite of the various views on the wave function in these theories, they never interpret the wave function as a description of the motion of particles in real space. However, on the one hand, the interpretation of the wave function in these theories is still an unsettled issue, and on the other hand, these theories may be not influenced by the interpretation in a significant way. Therefore, they may be consistent with the suggested interpretation of the wave function after certain revision.

4.1 The wavefunction collapse is probably real

4.1.1 No hidden variables

Let's first analyze the de Broglie-Bohm theory (de Broglie 1928; Bohm 1952). According to the theory, a complete realistic description of a quan-

[25]It has been argued that the wave function living on configuration space can hardly be considered as a real physical entity due to its multi-dimensionality (see, e.g. Monton 2002 and references therein). However, it seems that this common objection is not conclusive, and one can still insist on the reality of the wave function living on configuration space by resorting to some metaphysical arguments. For example, a general strategy is to show how a many-dimensional world can appear three-dimensional to its inhabitants, and then argue on that basis that a wavefunction ontology is adequate to explain our experience (Lewis 2004). As we have argued earlier, the existence of the effective mass and charge density of a quantum system, which is measurable by protective measurement, poses a more serious objection to the wavefunction ontology; even for a single quantum system the wave function cannot be taken as a field-like entity in three-dimensional space either. Moreover, the reason is not metaphysical but physical, i.e., that the field-like interpretation contradicts both quantum mechanics and experimental observations.

tum system is provided by the configuration defined by the position of its particle together with its wave function. The wave function follows the linear Schrödinger equation and never collapses. The particle, often called Bohmian particle, is guided by the wave function via the guiding equation to undergo continuous motion. The result of a measurement is indicated by the position of the Bohmian particle describing the pointer of the measuring device, and thus it is always definite. Moreover, it can be shown that the de Broglie-Bohm theory gives the same predictions of measurement results as standard quantum mechanics by means of a quantum equilibrium hypothesis (so long as the latter gives unambiguous predictions)[26]. In this way, it seems that the de Broglie-Bohm theory can succeed in avoiding the collapse of the wave function.

However, although the de Broglie-Bohm theory is mathematically equivalent to standard quantum mechanics, there is no clear consensus with regard to its physical interpretation. To begin with, the interpretation of the wave function in the theory is still in dispute. For example, the wave function has been regarded as a field similar to electromagnetic field (Bohm 1952), an active information field (Bohm and Hiley 1993), a field carrying energy and momentum (Holland 1993), a causal agent more abstract than ordinary fields (Valentini 1997), a component of physical law (Goldstein 2009), and a dispositional property of Bohmian particles (Belot 2011) etc. Notwithstanding the differences between these interpretations, they are inconsistent with the picture of random discontinuous motion of particles. To say the least, they can hardly explain the existence of mass and charge density for a charged quantum system, which is measurable by protective measurement. Our previous analysis shows that the mass and charge density of a quantum system, which is proportional to the modulus square of its wave function, is effective and formed by the ergodic motion of a localized particle with the total mass and charge of the system, and thus the wave function is a description of the ergodic motion of particles.

A more pivotal issue concerns the guiding responsibility of the wave function. In the de Broglie-Bohm theory, the wave function of a quantum system is assumed to guide the deterministic continuous motion of its Bohmian particle, while the wave function and the Bohmian particle are two different physical entities. According to our suggested interpretation of the wave function, the wave function of a quantum system indeed guides the motion of a localized particle in some sense. However, what the wave function describes is not a physical entity independent of the particle but an intrinsic property of the particle. Moreover, the motion of the particle is not

[26]Concretely speaking, the quantum equilibrium hypothesis provides the initial conditions for the guidance equation which make the de Broglie-Bohm theory obey Born's rule in terms of position distributions. Moreover, since all measurements can be finally expressed in terms of position, e.g. pointer positions, this amounts to full accordance with all predictions of quantum mechanics.

continuous and deterministic but discontinuous and random, and the wave function guides the discontinuous motion in a probabilistic way; the modulus square of the wave function determines the probability density of the particle appearing in every position in space. This reveals a deeper discrepancy between the de Broglie-Bohm theory and our suggested interpretation of the wave function.

Next, let's analyze the hypothetical Bohmian particles in the de Broglie-Bohm theory. It has been a controversial issue what properties the Bohmian particles should have. On the one hand, the theory seems to require that the Bohmian particles have mass and charge. For example, for a many-body system in an entangled state, the guiding equation of each Bohmian particle obviously contains the mass of each sub-system, and the mass is usually regarded as the mass of the Bohmian particle (Goldstein 2009). This attribution seems inevitable, as every sub-system does not have its own wave function, and thus the mass cannot be attributed to its wave function. Moreover, in the quantum potential formulation of the theory (Bohm 1952), the second-order equation of motion of the Bohmian particle of a charged quantum system contains a Coulomb force term when an electrostatic interaction is involved, which indicates that the Bohmian particle also has charge. Endowing mass and charge to the Bohmian particles seems quite natural, as for the de Broglie-Bohm theory the particles are primary or primitive, while the wave function is only secondary or derivative (Goldstein 2009).

On the other hand, it has been argued that the mass and charge of a quantum system should be possessed by its wave function, not by its Bohmian particle (see, e.g. Brown, Dewdney and Horton 1995). It is even claimed that a Bohmian particle has no properties other than its position (Hanson and Thoma 2011). As our previous analysis suggests, protective measurement may provide a more convincing argument for the "bareness" of the Bohmian particles. The existence of mass and charge density for a charged quantum system, which is proportional to the modulus square of its wave function and measurable by protective measurement, implies that mass and charge are attributes of the wave function and not attributes of the hypothetical Bohmian particle. When the wave function is further interpreted as a description of the random discontinuous motion of particles as we have suggested, it becomes more obvious that the mass and charge (and other properties) of a quantum system belong to these particles, not to the added Bohmian particles.

There is a possible way to avoid the above inconsistency. One may only use the first-order guiding equation to formulate the theory. There is no apparent appearance of charge in the equation even when an electrostatic interaction is involved; the charge information is absorbed into the wave function in some sense. Moreover, even if mass still appears in the guiding equation, one may attribute the masses of all sub-systems of a many-body system to its wave function. This seems to require a particle interpretation

of the wave function. For example, when interpreting the wave function as a description of the random discontinuous motion of particles, the masses appearing in the guiding equations can be attributed to these particles. In this way, it seems that the Bohmian particles can be consistently regarded as bare.

However, the "bareness" of the Bohmian particles is at least a worrisome issue, and it might already imply the non-existence of hidden variables at the worst. According to the common-sense view, a real particle should have its intrinsic properties such as mass and charge etc, and its total energy cannot be zero either. If a particle has no properties other than its position, then in what sense it can be regarded as physically real? It seems that a bare Bohmian particle has no difference with a mathematical point. Furthermore, if the Bohmian particles are deprived of all intrinsic properties, then how can they be guided by the wave function? and how can the wave function "know" its existence and guide its motion? This also reminds us another debatable aspect of the de Broglie-Bohm theory, the interaction between the wave function and the Bohmian particles. In the final analysis, the influence of the wave function on the Bohmian particles is in want of a *physical* explanation.

Lastly, we analyze a possible combining picture of the de Broglie-Bohm theory and our suggested interpretation of the wave function. Even if the wave function describes the random discontinuous motion of particles, it seems that one can also add the bare Bohmian particles undergoing continuous motion to explain the emergence of definite measurement results. This is one of the main merits of the de Broglie-Bohm theory after all. This hybrid theory, however, has more drawbacks. First of all, the double-particle picture seems clumsy and unnatural. For example, an ordinary electron will contain two distinct particles; one is a real localized electron, and the other is a bare particle without any properties of the electron. Moreover, these two particles move in two essentially different ways; the real electron undergoes random discontinuous motion, while the bare particle undergoes deterministic continuous motion. The coexistence of continuous motion and discontinuous motion seems inconsistent with the general expectation that motion can only be continuous or discontinuous in nature.

Secondly, the continuous motion of the bare particle needs to be guided by the real particle, but this added guiding responsibility can hardly be explained. It is natural that the wave function as an intrinsic property of the real particle determines the motion of the particle, but it seems difficult to explain why and how this property also determines the motion of another bare particle. Moreover, the determining ways are essentially different. The wave function guides the motion of the real particle in a probabilistic way, while it guides the motion of the bare particle in a deterministic way. Last but not the least, the trajectories of the bare particles as added hidden variables seems redundant. In some sense, there are already hidden variables

in the picture of random discontinuous motion of particles. They are the random position, momentum and energy of the particles at each instant. Though these variables are not continuous and deterministic, their random motion may just lead to the stochastic collapse of the wave function and further account for the emergence of definite and random measurement results. We will discuss this seemingly more natural possibility in detail later on.

In conclusion, when taking into account the implications of protective measurement and our suggested interpretation of the wave function based on them, the de Broglie-Bohm theory seems to be not a satisfactory solution to the measurement problem[27]. Although the theory can be mathematically equivalent to standard quantum mechanics, it seems lack of a reasonable physical interpretation. The added hidden variables, which are used to explain the emergence of definite measurement results, can only be carried by bare particles without any intrinsic properties of the involved quantum system such as mass and charge. Moreover, the theory can hardly explain why the evolution of the hidden variables is guided by the wave function in the way it requires. In particular, when the wave function is interpreted as a description of the random discontinuous motion of particles (and at a deeper level a description of the intrinsic property of the particles that determines their discontinuous motion in a probabilistic way), it seems impossible that the wave function belonging to these particles also guides the motion of other particles, especially when these particles are bare and the guiding way is deterministic.

4.1.2 No many worlds

Now let's turn to the second approach to avoid wavefunction collapse, the many-worlds interpretation. Although this theory is widely acknowledged as one of the main formulations of quantum mechanics, its many fundamental issues have not yet been solved (see Saunders et al 2010 and references therein). For example, the stuff of the many worlds, what they are made of, seems never adequately explained, nor are the worlds precisely defined. Moreover, no satisfactory role or substitute for probability has been found in the many worlds theories, and their consistency with quantum mechanics is still debatable. In the following, we will analyze whether there are many worlds according to the suggested interpretation of the wave function in terms of random discontinuous motion of particles.

In order to examine the validity of the many-worlds interpretation, it is crucial to know exactly what a quantum superposition is. No matter how to define the many worlds, they correspond to some branches of a quantum superposition after all (e.g. the branches where measuring devices

[27]This conclusion also applies to other hidden variables theories with added particle ontology.

obtain definite results, and in particular, observers have definite conscious experience). According to the picture of random discontinuous motion of particles, a quantum superposition exists in the form of random and discontinuous time division. For a superposition of two positions A and B of a quantum system, the system randomly and discontinuously moves between these two positions. At some random and discontinuous instants the system is in position A, and at other instants it is in position B[28]. As a result, each position branch exists in a time sub-flow, and the sum of all these time sub-flows constitute the whole continuous time flow. In this picture of quantum superposition, it is obvious that there is only one system all along in the continuous time flow, which randomly and discontinuously moves throughout all branches of the superposition, no matter the system is a microscopic particle or a measuring device or an observer. In other words, there is only one world which instantaneous state is constantly changing in a random and discontinuous way.

This conclusion is also supported by a comparison between discontinuous motion and continuous motion. For a quantum particle undergoing discontinuous motion, the position of the particle changes discontinuously. For a classical particle, its position changes continuously. There is no essential difference between these two kinds of changes. For both cases the position of the particle is always definite at each instant, and the positions of the particle at different instants may be different. Moreover, the discontinuous change, like the continuous change, does not result in the process needed for creating the many worlds, because, among other reasons, the change happens all the while but the creating process only happens once. Therefore, if there is only one world in classical mechanics, so does in quantum mechanics according to the picture of random discontinuous motion of particles, no matter how the many worlds are defined.

To sum up, the above analysis indicates that the de Broglie-Bohm theory and the many-worlds interpretation are not satisfactory solutions to the measurement problem according to our suggested interpretation of the wave function in terms of random discontinuous motion of particles. If there are neither hidden variables nor many worlds that can explain the emergence of definite measurement results, then the collapse of the wave function is probably a real physical process, which is responsible for the transition from microscopic uncertainty to macroscopic certainty. Accordingly, the dynamical collapse theories may be in the right direction by admitting wavefunction collapse (Ghirardi 2008).

As noted earlier, however, the existing ontology of the dynamical collapse theories, such as mass density ontology and flashes ontology (Ghirardi, Grassi and Benatti 1995; Ghirardi 1997, 2008; Allori et al 2008), is incon-

[28]That the system is in a definite position A or B at every instant already implies that there is only one world at any time.

sistent with our suggested interpretation of the wave function. Especially, the existence of the effective mass and charge density of a quantum system, which is measurable by protective measurement, seems to already exclude the mass density ontology. In addition, the existing dynamical collapse theories are still phenomenological models, and they are also plagued by some serious problems such as energy non-conservation etc (Pearle 2007, 2009). In particular, the physical origin of the wavefunction collapse, including the origin of the randomness of the collapse process, is still unknown, though there are already some interesting conjectures (see, e.g. Diósi 1989; Penrose 1996). In the following subsections, we will try to solve these problems and propose a new dynamical collapse model in terms of the random discontinuous motion of particles. A more detailed review of the existing dynamical collapse theories will be given in the last subsection.

4.2 A conjecture on the origin of wavefunction collapse

It is well known that a chooser and choices are needed to bring the required dynamical collapse of the wave function (Pearle 1999). The chooser is the noise source that collapses the wave function, and the choices are the states toward which the collapse tends. In this section, we will first analyze these two relatively easier problems and then investigate the more difficult problem, the physical origin of wavefunction collapse.

4.2.1 The chooser in discrete time

To begin with, let's analyze the chooser problem. In the existing dynamical collapse models, the chooser is generally assumed to be an unknown classical noise field independent of the collapsed wave function (Pearle 2007, 2009). If what the wave function describes is the random discontinuous motion of particles, then it seems natural to assume that the random motion of particles is the appropriate noise source to collapse the wave function. This has three merits at least. First, the noice source and its properties are already known. For example, the probability of the particles being in certain position, momentum and energy at each instant is given by the modulus square of their wave function at the instant. Next, this noice source is not a classical field, and thus the model can avoid the problems introduced by the field such as its quantization etc. Last but not the least, the random discontinuous motion of particles can also manifest itself in the equation of motion by introducing the collapse evolution of the wave function. In the following, we will give a more detailed argument for this assumption.

According to our suggested interpretation of the wave function, the wave function of a quantum particle can be regarded as an instantaneous intrinsic property of the particle that determines its random discontinuous motion. However, the wave function is not a complete description of the instanta-

neous state of the particle. The instantaneous state of the particle at a given instant also includes its random position, momentum and energy at the instant, which may be called the random part of the instantaneous state of the particle. Although the probability of the particle being in each random instantaneous state is completely determined by the wave function, its stay in the state at each instant is a new physical fact independent of the wave function. Therefore, it seems natural to assume that the random stays of the particle may have certain physical efficiency that manifests in the complete equation of motion [29]. Since the motion of the particle is essentially random, its stay at an instant does not influence its stays at other instants in any direct way. Then the random stays of the particle can only manifest themselves in the equation of motion by their influences on the evolution of the wave function[30]. This forms a feedback in some sense; the wave function of a particle determines the probabilities of its stays in certain position, momentum and energy, while its random stay at each instant also influences the evolution of the wave function in a stochastic way[31].

However, the existence of the influences of the random motion of a particle relies on an important precondition, the discreteness of time. If time is continuous and instants are durationless, the random stays of a particle can have no influence on anything. The reason is as follows. First, the duration of each random stay of the particle is zero in continuous time. Due to the *randomness* of motion, when there are at least two possible instantaneous states a particle can move throughout, the particle cannot stay in the same instantaneous state all through for a finite time. For the joint probability of the particle being in the same instantaneous state for all infinitely uncountable instants in the finite time interval is obviously zero, and the total probability of the particle being in other instantaneous states is not zero at any instant in between either. In other words, in order that a particle stays in the same instantaneous state for a finite time, the probability of the particle being in this instantaneous state must be one all the while during the entire interval. This is possible only for the banal case where there is only one instantaneous state the particle can stay and thus there is no motion and its randomness at all throughout the duration[32]

[29]This is distinct from the case of continuous motion. For the latter, the position of a particle at each instant is *completely* determined by the deterministic instantaneous condition at the instant, and thus the position of the particle has no influence on the deterministic instantaneous condition.

[30]In fact, since the random stays of a particle as one part of its instantaneous state are completely random, the complete evolution equation of the instantaneous state of the particle is only about the evolution of the wave function. Therefore, the random stays of the particle can only manifest themselves in the complete equation of motion by their stochastic influences on the evolution of the wave function.

[31]In other words, the wave function of a particle determines its random discontinuous motion, while the motion also influences the evolution of the wave function reciprocally.

[32]Unfortunately, this banal case does not exist. Due to the uncertainty relation between

Secondly, the influence of the random stay of a particle at a durationless instant is zero, and the influence cannot accumulate from zero to a finite value during a time interval either due to the *discontinuity* of motion. This is distinct from the situation of continuous motion. The influence of a continuous process can accumulate from zero at each instant to a finite value during a time interval. Another way to understand this result is to see that the discontinuity and randomness of motion exist only at each instant, and they don't exist during a finite time interval or even an infinitesimal time interval. For example, the state of random discontinuous motion in real space, which is defined during an infinitesimal time interval, is described by the position density and position flux density, and they are continuous quantities that contains no discontinuity and randomness.

Therefore, if time is continuous and instants are durationless, then the random stays of a particle can have no physical effect even during a finite time interval. This also means that the random stays of a particle can influence the evolution of its wave function only when instants are not zero-sized but finite-sized, i.e., when time is essentially discrete or quantized. Once the duration of each random stay of a particle is finite, each random stay can have a finite influence on the evolution of the wave function. It is worth stressing again that if time is not discrete but continuous, a particle cannot stay in one of the infinitely many instantaneous states all through for a finite time; rather, it can only stay there for one zero-sized instant. But if time is discrete and instants are not zero-sized but finite-sized, even if a particle stays in an instantaneous state only for one instant, the duration of its stay is also finite as the instant is finite-sized. In some sense, the discreteness of time prevents a particle from jumping from its present instantaneous state to another instantaneous state and makes the particle stay in the present instantaneous state all through during each finite-sized instant. Since it has been widely conjectured that spacetime is discrete and the Planck scale is the minimum spacetime scale[33], we will assume that the size of each discrete instant or the quantum of time is the Planck time in our following analysis.

To sum up, the manifestability of the randomness and discontinuity of motion in the laws of motion requires that time is essentially discrete. In the discrete time, a particle randomly stays in an instantaneous state with definite position, momentum and energy at each discrete instant, which probability is determined by the modulus square of its wave function at the instant, and each random, finite stay of the particle may have a finite

position and momentum in quantum mechanics, there are always infinitely many different instantaneous states (with definite position and momentum) where a particle can stay at any time.

[33]Note that the existing arguments, which are based on some sort of combination of quantum theory and general relativity (see, e.g. Garay 1995 for a review), do not imply but only suggest that spacetime is discrete. Moreover, the meanings and realization of discrete spacetime are also different in the existing models of quantum gravity.

influence on the evolution of its wave function. As we will show later, the accumulation of such discrete and random influences may lead to the right collapse of the wave function, which can explain the emergence of definite measurement results. Accordingly, the evolution of the wave function will be governed by a revised Schrödinger equation, which includes the normal linear terms and a stochastic nonlinear term that describes the discrete collapse dynamics. Note that the wave function (as intrinsic property of particles) also exists in the discrete time, which means that the wave function does not change during each discrete instant, and the evolution of the wave function including the linear Schrödinger evolution is also essentially discrete.

4.2.2 Energy conservation and the choices

Now let's investigate the choice problem. The random stay of a particle may have a stochastic influence on the evolution of its wave function at each discrete instant. Then when the stochastic influences accumulate and result in the collapse of the wave function, what are the states toward which collapse tends? This is the choice problem or preferred basis problem. It may be expected that the stochastic influences of the motion of a particle on its wave function should not be arbitrary but be restricted by some fundamental principles. In particular, it seems reasonable to assume that the resulting dynamical collapse of the wave function should also satisfy the conservation of energy. As a result, the collapse states or choices will be the energy eigenstates of the total Hamiltonian of the system[34]. In the following, we will give a more detailed analysis of the consequences of this assumption. Its possible physical basis will be investigated in the next subsection.

As we have argued in the last section, for a deterministic evolution of the wave function such as the linear Schrödinger evolution, the requirement of energy conservation applies to a single isolated system. However, for a stochastic evolution of the wave function such as the dynamical collapse process, the requirement of energy conservation cannot apply to a single system in general but apply to an ensemble of identical systems[35]. It can be proved that only the collapse states are energy eigenstates of the total Hamiltonian of a given system, can energy be conserved for an ensemble of identical systems for wavefunction collapse (See Pearle 2000 for a more detailed analysis). Note that for the linear Schrödinger evolution under an external potential, energy is conserved but momentum is not conserved even at the ensemble level, and thus it is not momentum conservation but energy conservation that is a more universal restriction for wavefunction collapse.

The conservation of energy can not only help to solve the preferred basis

[34]For the superpositions of degenerate energy eigenstates of a many-particle system, a further collapse rule is needed. We will discuss this issue later on.

[35]As we will see later, the conservation of energy may also hold true at the individual level for the collapse evolution of some special wave functions.

problem, but also further determine the law of dynamical collapse to a large extent. For each system in the same quantum state in an ensemble, in order that the probability distribution of energy eigenvalues of the state can keep constant for the whole ensemble (i.e. energy is conserved at the ensemble level), the random stay of the system at each discrete instant can only change its (objective) energy probability distribution[36], and moreover, the change must also satisfy a certain restriction. Concretely speaking, the random stay in a definite energy E_i will increase the probability of the energy branch $|E_i>$ and decrease the probabilities of all other energy branches pro rata. Moreover, the increasing amplitude must be proportional to the total probability of all other energy branches, and the coefficient is related to the energy uncertainty of the state. We will demonstrate this result in the next subsection.

A more important problem is whether this energy-conserved collapse model can explain the emergence of definite measurement results and our macroscopic experience. At first sight the answer appears negative. For example, the energy eigenstates being collapse states seems apparently inconsistent with the localization of macroscopic objects. However, a detailed analysis given in the following subsections will demonstrate that the model can be consistent with existing experiments and our macroscopic experience. The key point is that the energy uncertainty driving the collapse of the entangled state of a many-body system is not the uncertainty of the total energy of all sub-systems, but the sum of the absolute energy uncertainty of every sub-system. As a result, the collapse states are the product states of the energy eigenstates of the Hamiltonian of each sub-system for a non-interacting or weakly-interacting many-body system. This provides a further collapse rule for the superpositions of degenerate energy eigenstates of a many-body system.

4.2.3 In search of a deeper basis

In this subsection, we will investigate the possible physical basis of the energy conservation restriction for wavefunction collapse.

It is well known that the conservation of energy and momentum refers to an ensemble of identical systems in standard quantum mechanics. However, this standard view is based on the probability interpretation of the wave function, and it seems unnatural when assuming an objective interpretation of the wave function of a single system, e.g. our suggested interpretation in terms of random discontinuous motion of particles. An ensemble is not an actual system after all, and the conservation of something for an emsemble seems meaningless. Moreover, since a single system in the ensemble does not

[36]If the phase of an energy eigenstate also changes with time, then the probability distribution of energy eigenvalues will in general be changed for each identical system in the ensemble, and as a result, energy will be not conserved even at the ensemble level.

know other systems and the whole ensemble, there must exist some underlying mechanism that can ensure the conservation of energy for an ensemble. Then the conservation of energy for an ensemble of identical systems is probably a result of the laws of motion for individual systems in the ensemble. Here is a possible scheme. First of all, energy is conserved for the evolution of individual energy eigenstates. Next, a superposition of energy eigenstates will dynamically collapse to one of these energy eigenstates, and the probability of the collapse result satisfies Born's rule. Then the wavefunction collapse will satisfy the conservation of energy for an ensemble of identical systems.

In the following, we will further suggest a possible physical basis for this scheme of energy-conserved wavefunction collapse. According to the picture of random discontinuous motion, for a particle in a superposition of energy eigenstates, the particle stays in an instantaneous state with a definite energy eigenvalue at a discrete instant, and at another instant it may jump to another instantaneous state with another energy eigenvalue. It seems to be a reasonable assumption that the particle has both the tendency to jump among the instantaneous states with different energies and the tendency to stay in the instantaneous states with the same energy, and their relative strength is determined by the energy probability distribution of the particle. This is satisfactory in logic, as there should exist two opposite tendencies in general, and their relative strength is determined by certain condition. In some sense, the two tendencies of a particle are related to the two parts of its instantaneous state; the jumping tendency is related to the wave function, and it is needed to manifest the superposition of different energy eigenstates, while the staying tendency is related to the random stays. These two opposite tendencies together constitute the complete "temperament" of a particle.

It can be argued that the tendency to stay in the same energy for individual particles might be the physical origin of the energy-conserved wavefunction collapse. For a particle in a superposition of energy eigenstates, the particle stays in an instantaneous state with definite energy at a discrete instant, and the staying tendency of the particle will increase its probability of being in the instantaneous states with the present energy at next instant. In other words, the random stay of a particle in an instantaneous state with an energy eigenvalue will increase the probability of the energy eigenvalue (and correspondingly decrease the probabilities of other energy eigenvalues pro rata). Moreover, the increase of probability may relate to the energy probability distribution of the particle. By the continuity of change of staying tendency, the particle will jump more readily among the instantaneous states with small energy uncertainty and more hardly among the instantaneous states with large energy uncertainty (which can also be regarded as a restriction of energy change). Thus the larger the energy uncertainty of the superposition is, the larger the increase of probability is during each

random stay. A detailed calculation, which will be given in the next subsection, shows that such random changes of energy probability distribution can continuously accumulate to lead to the collapse of the superposition of energy eigenstates to one of them.

It can be further argued that the probability distribution of energy eigenvalues should keep constant during the random evolution of an ensemble of identical systems, and thus the resulting wavefunction collapse will satisfy Born's rule. The reason is as follows. When an initial superposition of energy eigenstates undergoes the dynamical collapse process, the probability distribution of energy eigenvalues should manifest itself through the collapse results for an ensemble of identical systems. At a deeper level, it is very likely that the laws of nature permit nature to manifest itself as best as one can, or else we will be unable to find the laws of nature and verify them by experiments, and our scientific investigations will be also pointless. This may be regarded as a law of laws. Since the collapse evolution of individual systems is competely random and irreversible, the diagonal density matrix elements for an ensemble of identical systems must be precisely the same as the initial probability distribution at every step of the evolution. Otherwise the frequency distribution of the collapse results in the ensemble cannot reflect the initial probability distribution, or in other words, the probability information contained in the initial state will be completely lost due to the random and irreversible collapse[37]. As a consequence, the collapse evolution will conserve energy at the ensemble level, and the collapse results will also satisfy Born's rule in quantum mechanics.

Certainly, there is still a question that needs to be answered. Why energy? Why not position or momentum? If there is only one property that undergoes the random discontinuous motion (e.g. position), then the above tendency argument for the unique property may be satisfying. But if there are many properties that undergoes the random discontinuous motion, then we need to answer why the tendency argument applies only to energy. A possible answer is that energy is the property that determines the linear evolution of the state of motion, and thus it seems natural and uniform that energy also determines the nonlinear collapse evolution. Moreover, energy eigenstates are the states of motion that no longer evolve (except an absolute phase) for the linear evolution. Then by analogy, it is likely that energy eigenstates are also the states that no longer evolve for the nonlinear collapse evolution, i.e., that the energy eigenstates are the collapse states. However, we may never be able to reach (and know we reach) the end point of explanation. Another important thing is to develop a concrete model and compare it with experiments. Let's do this in the following subsections.

[37]Note that the reversible Schrödinger evolution conserves the information even for individual isolated systems.

4.3 A discrete model of energy-conserved wavefunction collapse

After giving a very speculative analysis of the origin of wavefunction collapse in terms of random discontinuous motion of particles, we will propose a discrete model of energy-conserved wavefunction collapse based on some results obtained from the analysis.

Consider a multi-level system with a constant Hamiltonian. Its initial state is:

$$|\psi(0)\rangle = \sum_{i=1}^{m} c_i(0)\,|E_i\rangle, \tag{48}$$

where $|E_i\rangle$ is the energy eigenstate of the Hamiltonian of the system, E_i is the corresponding energy eigenvalue, and $c_i(0)$ satisfies the normalization relation $\sum_{i=1}^{m} |c_i(0)|^2 = 1$.

According to our conjecture on the origin of wavefunction collapse, this superposition of energy eigenstates will collapse to one of the eigenstates after a discrete dynamical process, and the collapse evolution satisfies the conservation of energy at the ensemble level. The physical picture of the dynamical collapse process is as follows. At the initial discrete instant $t = t_P$ (where t_P is the Planck time), the system randomly stays in a branch $|E_i\rangle$ with probability $P_i(0) \equiv |c_i(0)|^2$[38]. This finite stay slightly increases the probability of the staying branch and decreases the probabilities of all other branches pro rata. Similarly, at any discrete instant $t = nt_P$ the system randomly stays in a branch $|E_i\rangle$ with probability $P_i(t) \equiv |c_i(t)|^2$, and the random stay also changes the probabilities of the branches slightly. Then during a finite time interval much larger than t_P, the probability of each branch will undergo a discrete and stochastic evolution. In the end, the probability of one branch will be close to one, and the probabilities of other branches will be close to zero. In other words, the initial superposition will randomly collapse to one of the energy branches in the superposition.

Now we will give a concrete analysis of this dynamical collapse process. Since the linear Schrödinger evolution does not change the energy probability distribution, we may only consider the influence of dynamical collapse on the energy probability distribution. Suppose the system stays in branch $|E_i\rangle$ at the discrete instant $t = nt_P$, and the stay changes the probability of this branch, $P_i(t)$, to

$$P_i^i(t + t_P) = P_i(t) + \Delta P_i, \tag{49}$$

[38]Strictly speaking, the description "branch" should be replaced by "instantaneous state", e.g. the branch $|E_i\rangle$ should be replaced by the instantaneous state with energy E_i. Yet the branch description may be more succinct and visual, and we will use it in the following discussions.

where the superscript i denotes the staying branch, and ΔP_i is a functional of $P_i(t)$. Due to the conservation of probability, the increase of the probability of one branch can only come from the scale-down of the probabilities of all other branches. This means that the probablity of another branch $P_j(t)$ $(j \neq i)$ correspondingly turns to be[39]

$$P_j^i(t+t_P) = P_j(t) - \frac{P_j(t)\Delta P_i}{1-P_i(t)}, \tag{50}$$

where the superscript i also denotes the staying branch. The probability of this random stay at the instant is $p(E_i,t) = P_i(t)$. Then we can work out the diagonal density matrix elements of the evolution[40]:

$$\begin{aligned}
\rho_{ii}(t+t_P) &= \sum_{j=1}^{m} p(E_j,t)P_i^j(t+t_P) \\
&= P_i(t)[P_i(t)+\Delta P_i] + \sum_{j\neq i} P_j(t)[P_i(t) - \frac{P_i(t)\Delta P_j(t)}{1-P_j(t)}] \\
&= \rho_{ii}(t) + P_i(t)[\Delta P_i - \sum_{j\neq i} P_j(t)\frac{\Delta P_j(t)}{1-P_j(t)}].
\end{aligned} \tag{51}$$

Here we shall introduce the first rule of dynamical collapse, which says that the probability distribution of energy eigenvalues for an ensemble of identical systems is constant during the dynamical collapse process. As we have argued in the last subsection, this rule is required by the principle of energy conservation at the ensemble level, and it may also have a physical basis relating to the manifestability of nature. By this rule, we have $\rho_{ii}(t+t_P) = \rho_{ii}(t)$ for any i. This leads to the following equations set:

$$\begin{aligned}
\Delta P_1(t) - \sum_{j\neq 1} \frac{P_j(t)\Delta P_j(t)}{1-P_j(t)} &= 0, \\
\Delta P_2(t) - \sum_{j\neq 2} \frac{P_j(t)\Delta P_j(t)}{1-P_j(t)} &= 0, \\
&\cdots \\
\Delta P_m(t) - \sum_{j\neq m} \frac{P_j(t)\Delta P_j(t)}{1-P_j(t)} &= 0.
\end{aligned} \tag{52}$$

[39] One can also obtain this result by first increasing the probability of the staying branch and then normalizing the probabilities of all branches. This means that $P_i(t+t_P) = \frac{P_i(t)+\Delta}{1+\Delta}$ and $P_j(t+t_P) = \frac{P_j(t)}{1+\Delta}$ for any $j \neq i$. In this way, we have $\Delta P_i = \frac{\Delta}{1+\Delta}(1-P_i(t))$ and $\Delta P_j = \frac{\Delta}{1+\Delta}P_j(t)$ for any $j \neq i$.

[40] The density matrix describes the ensemble of states which arise from all possible random stays.

By solving this equations set (e.g. by subtracting each other), we find the following relation for any i:

$$\frac{\Delta P_i}{1 - P_i(t)} = k, \tag{53}$$

where k is an undetermined dimensionless quantity that relates to the state $|\psi(t)\rangle$.

By using Eq. (53), we can further work out the non-diagonal density matrix elements of the evolution. But it is more convenient to calculate the following variant of non-diagonal density matrix elements:

$$\begin{aligned}
\varrho_{ij}(t + t_P) &= \sum_{l=1}^{m} p(E_l, t) P_i^l(t + t_P) P_j^l(t + t_P) \\
&= \sum_{l \neq i,j} P_l(t)[P_i(t) - kP_i(t)][P_j(t) - kP_j(t)] \\
&\quad + P_i(t)[P_i(t) + k(1 - P_i(t))][P_j(t) - kP_j(t)] \\
&\quad + P_j(t)[P_j(t) + k(1 - P_j(t))][P_i(t) - kP_i(t)] \\
&= (1 - k^2)\varrho_{ij}(t).
\end{aligned} \tag{54}$$

Since the usual collapse time, τ_c, is defined by the relation $\varrho_{ij}(\tau_c) = \frac{1}{2}\varrho_{ij}(0)$, we may use a proper approximation, where k is assumed to be the same as its initial value during the time interval $[0, \tau_c]$, to simplify the calculation of the collapse time. Then we have:

$$\varrho_{ij}(t) \approx (1 - k^2)^n \varrho_{ij}(0). \tag{55}$$

The corresponding collapse time is in the order of:

$$\tau_c \approx \frac{1}{k^2} t_P, \tag{56}$$

In the following, we shall analyze the formula of k defined by Eq. (53). To begin with, the probability restricting condition $0 \leqslant P_i(t) \leqslant 1$ for any i requires that $0 \leqslant k \leqslant 1$. When $k = 0$, no collapse happens, and when $k = 1$, collapse happens instantaneously. Note that k cannot be smaller than zero, as this will lead to the negative value of $P_i(t)$ in some cases. For instance, when k is negative and $P_i(t) < \frac{|k|}{1+|k|}$, $P_i(t + t_P) = P_i(t) + k[1 - P_i(t)]$ will be negative and violate the probability restricting condition. That k is positive indicates that each random stay increases the probability of the staying branch and decreases the probabilities of other branches, which is consistent with the analysis given in the last subsection.

Next, k is proportional to the duration of stay. The influence of each stay on the probability of the staying branch is an accumulating process. When the duration of stay is zero as in continuous space and time, no influence

exists and no collapse happens. When the duration of stay, t_P, is longer, the probability of the staying branch will increase more. Thus we have $k \propto t_P$.

Thirdly, k is also proportional to the energy uncertainty of the superposition of energy eigenstates. First, from a dimensional analysis k should be proportional to an energy term in order to cancel out the dimension of time. Next, the energy term should be the energy uncertainty of the superposition defined in an appropriate way according to the analysis of the last subsection. When the energy uncertainty is zero, i.e., when the state is an energy eigenstate, no collapse happens. When the energy uncertainty is not zero, collapse happens. Moreover, the larger the energy uncertainty is, the larger the increase of the probability of the staying branch for each random stay is, namely the larger k is. Therefore, k will be proportional to the energy uncertainty of the superposition. How to define the energy uncertainty then? Since k is invariant under the swap of any two branches (P_i, E_i) and (P_j, E_j) according to Eq. (53), the most natural definition of the energy uncertainty of a superposition of energy eigenstates is[41]:

$$\Delta E = \frac{1}{2} \sum_{i,j=1}^{m} P_i P_j |E_i - E_j|. \tag{57}$$

It seems a little counterintuitive that k contains the energy uncertainty term that relates to the whole energy distribution. The puzzle is two-fold. First, this means that the increase of the probablity of the staying branch relates not to the energy difference between the staying branch and all other branches, but to the energy uncertainty of the whole state. This is reflected in the formula of ΔE in the existence of the energy difference between any two branches, $|E_i - E_j|$ for any i and j. Next, the increase of the probablity of the staying branch relates also to the energy probablity distribution that determines the energy uncertainty. This is reflected in the formula of ΔE in the existence of $P_i P_j$. In fact, these seemingly puzzling aspects are still understandable. The first feature is required by the first rule of dynamical collapse that ensures energy conservation at the ensemble level. This can be clearly seen from Eq. (53). If the increase of the probablity of the staying branch relates to the difference between the energy of the staying branch and the average energy of all other branches, then Eq. (53) will not hold true because the swap symmetry of k will be violated, and as a result, the first rule of dynamical collapse will be broken. The second feature can be understood as follows. In the picture of random discontinuous motion, the probability distribution contains the information of staying time distribution. An energy branch with small probability means that the system jumps through it less frequently. Thus this energy branch only makes a small contribution to the restriction of energy change or the increase of the staying tendency. As a

[41] Note that the common RMS (mean square root) uncertainty also satisfies the swap symmetry.

result, k or the increase of the probablity of the staying branch will relate not only to energy difference, but also to the energy probablity distribution.

Then after omitting a coefficient in the order of unity, we can get the formula of k in the first order:

$$k \approx \Delta E t_P/\hbar. \tag{58}$$

This is the second rule of dynamical collapse. By inputing Eq. (58) into Eq. (56), we can further get the collapse time formula:

$$\tau_c \approx \frac{\hbar E_P}{(\Delta E)^2}, \tag{59}$$

where $E_P = h/t_P$ is the Planck energy, and ΔE is the energy uncertainty of the initial state.

Here it is worth pointing out that k must contain the first order term ΔE. For the second order term $(\Delta E)^2$ or higher order terms will lead to much longer collapse time for some common measurement situations, which contradicts experiments (Gao 2006). Besides, a similar analysis of the consistency with experiments may also provide a further support for the energy-conserved collapse model in which the collapse states are energy eigenstates. First of all, if the collapse states are not energy eigenstates but momentum eigenstates, then the energy uncertainty will be replaced by momentum uncertainty in the collapse time formula Eq. (59), and thus the formula becomes $\tau_c \approx \frac{\hbar E_P}{(\Delta pc)^2}$. As a result, the collapse time will be too short to be consistent with experiments for some situations. For example, the collapse time will be about several days for the ground state of a hydrogen atom. Note that the second order term $(\Delta p)^2$ will also lead to much longer collapse time for some common measurement situations, which contradicts experiments.

Next, if the collapse states are position eigenstates[42], then the collapse time formula Eq. (59) will be replaced by something like $\tau_c \approx \frac{l^2 t_P}{(\Delta x)^2}$, where l is a finite length scale relating to the studied system. No matter what length scale l is, the collapse time of a momentum eigenstate will be zero as its position uncertainty is infinite. This means that the momentum eigenstates of any quantum system including microscopic particles will collapse instantaneously to one of its position eigenstates and thus cannot exist. Moreover, the superposition states with very small momentum uncertainty will also collapse very quickly even for microscopic particles. These results are apparently inconsistent with quantum mechanics. Although it may be possible to adjust the length scale l to make the model consistent with experiments, the collapse time formula will be much more complex than that

[42]In continuous space and time, a position eigenstate has infinite average energy and cannot be physically real. But in discrete space and time, position eigenstates will be the states which spatial dimension is about the Planck length, and they may exist.

of the energy-conserved collapse model. Let's give a little more detailed analysis here. There are two universal length scales for a quantum system: its Compton wavelength λ_c and the Planck length l_P. It is obvious that both of them cannot be directly used as the length scale in the above collapse time formula. Then the formula can only be written in a more complex form: $\tau_c \approx (\frac{\lambda_c}{l_P})^\alpha \cdot \frac{{\lambda_c}^2 t_P}{(\Delta x)^2}$. Moreover, the existing experiments such as the SQUID experiments and our everyday macroscopic experience require $\alpha \approx 8$. It seems very difficult to explain this unusually large exponent in a fundamental theory like the collapse dynamics. To sum up, the collapse states can hardly be position eigenstates when considering the consistency with experiments and the simplicity of theory.

Based on the above analysis, the state of the multi-level system at instant $t = nt_P$ will be:

$$|\psi(t)\rangle = \sum_{i=1}^{m} c_i(t) e^{-iE_i t/\hbar} |E_i\rangle, \tag{60}$$

Besides the linear Schrödinger evolution, the collapse dynamics adds a discrete stochastic evolution for $P_i(t) \equiv |c_i(t)|^2$:

$$P_i(t + t_P) = P_i(t) + \frac{\Delta E}{E_P}[\delta_{E_s E_i} - P_i(t)], \tag{61}$$

where ΔE is the energy uncertainty of the state at instant t defined by Eq. (57), E_s is a random variable representing the random stay of the system, and its probability of assuming E_i at instant t is $P_i(t)$. When $E_s = E_i$, $\delta_{E_s E_i} = 1$, and when $E_s \neq E_i$, $\delta_{E_s E_i} = 0$.

This equation of dynamical collapse can be directly extended to the entangled states of a many-body system. The difference only lies in the definition of the energy uncertainty ΔE. According to our analysis in the last subsection, for a non-interacting or weakly-interacting many-body system in an entangled state, for which the energy uncertainty of each sub-system can be properly defined, ΔE is the sum of the absolute energy uncertainty of all sub-systems, namely

$$\Delta E = \frac{1}{2} \sum_{l=1}^{n} \sum_{i,j=1}^{m} P_i P_j |E_{li} - E_{lj}|, \tag{62}$$

where n is the total number of the entangled sub-systems, m is the total number of energy branches in the entangled state, and E_{li} is the energy of sub-system l in the i-th energy branch of the state. Correspondingly, the collapse states are the product states of the energy eigenstates of the Hamiltonian of each sub-system. It should be stressed here that ΔE is not defined as the uncertainty of the total energy of all sub-systems as in the energy-driven collapse models (see, e.g. Percival 1995, 1998; Hughston

1996). For each sub-system has its own energy uncertainty that drives its collapse, and the total driving "force" for the whole entangled state should be the sum of the driving "forces" of all sub-systems, at least in the first order approximation. Although these two kinds of energy uncertainty are equal in numerical values in some cases (e.g. for a strongly-interacting many-body system), there are also some cases where they are not equal. For example, for a superposition of degenerate energy eigenstates of a non-interacting many-body system, which may arise during a common measurement process, the uncertainty of the total energy of all sub-systems is exactly zero, but the absolute energy uncertainty of each sub-system and their sum may be not zero. As a result, the superpositions of degenerate energy eigenstates of a many-particle system may also collapse. As we will see later, this is an important feature of our model, which can avoid Pearle's (2004) serious objections to the energy-driven collapse models.

It can be seen that the equation of dynamical collapse, Eq.(61), has an interesting property, scale invariance. After one discrete instant t_P, the probability increase of the staying branch $|E_i\rangle$ is $\Delta P_i = \frac{\Delta E}{E_P}(1-P_i)$, and the probability decrease of the neighboring branch $|E_{i+1}\rangle$ is $\Delta P_{i+1} = \frac{\Delta E}{E_P}P_{i+1}$. Then the probability increase of these two branches is

$$\Delta(P_i + P_{i+1}) = \frac{\Delta E}{E_P}[1 - (P_i + P_{i+1})]. \tag{63}$$

Similarly, the equation $\Delta P = \frac{\Delta E}{E_P}(1-P)$ holds true for the total probability of arbitrarily many branches (one of which is the staying branch). This property of scale invariance may simplify the analysis in many cases. For example, for a superposition of two wavepackets with energy difference, ΔE_{12}, much larger than the energy uncertainty of each wavepacket, $\Delta E_1 = \Delta E_2$, we can calculate the collapse dynamics in two steps. First, we use Eq.(61) and Eq.(57) with $|E_1 - E_2| = \Delta E_{12}$ to calculate the time of the superposition collapsing into one of the two wavepackets[43]. Here we need not to consider the almost infinitely many energy eigenstates constituting each wavepacket and their probability distribution. Next, we use Eq.(61) with $\Delta E = \Delta E_1$ to calculate the time of the wavepacket collapsing into one of its energy eigenstates. In general, this collapse process is so slow that its effect can be ignored.

Lastly, we want to stress another important point. As we have argued before, the discontinuity of motion requires that the collapse dynamics must be discrete in nature, and moreover, the collapse states must be energy eigenstates in order that the collaspe dynamics satisfies the conservation of energy at the ensemble level. As a result, the energy eigenstates and their corresponding eigenvalues must be also discrete for any quantum system.

[43]Note that most collapse states in an ensemble of identical systems keep the shape of the wavepacket almost precisely.

This result seems to contradict quantum mechanics, but when considering that our universe has a finite size (i.e. a finite event horizon), the momentum and energy eigenvalues of any quantum system in the universe may be indeed discrete[44]. The reason is that all quantum systems in the universe are limited by the finite horizon, and thus no free quantum systems exist in the strict sense. For example, the energy of a massless particle (e.g. photon) can only assume discrete values $E_n = n^2 \frac{hc}{4R_U}$, and the minimum energy is $E_1 = \frac{hc}{4R_U} \approx 10^{-33} eV$, where $R_U \approx 10^{25} m$ is the radius of the horizon of our universe. Besides, for a free particle with mass m_0, its energy also assumes discrete values $E_n = n^2 \frac{h^2}{32 m_0 R_U^2}$. For instance, the minimum energy is $E_1 \approx 10^{-72} eV$ for free electrons, which is much smaller than the minimum energy of photons.

It is interesting to see whether this tiny discreteness of energy makes the collapse dynamics more abrupt. Suppose the energy uncertainty of a quantum state is $\Delta E \approx 1eV$, and its energy ranges between the minimum energy E_1 and $1eV$. Then we can get the maximum energy level $l_{max} \approx \sqrt{\frac{1eV}{10^{-33}eV}} \approx 10^{16}$. The probability of most energy eigenstates in the superposition will be about $P \approx 10^{-16}$. During each discrete instant t_P, the probability increase of the staying energy branch is $\Delta P \approx \frac{\Delta E}{E_P}(1 - P) \approx 10^{-28}$. This indicates that the probability change during each random stay is still very tiny. Only when the energy uncertainty is larger than $10^{23} eV$ or $10^{-5} E_P$, will the probability change during each random stay be sharp. Therefore, the collapse evolution is still very smooth for the quantum states with energy uncertainty much smaller than the Planck energy.

4.4 On the consistency of the model and experiments

In this section, we will analyze whether the discrete model of energy-conserved wavefunction collapse is consistent with existing experiments and our macroscopic experience. Note that Adler (2002) has already presented a detailed consistency analysis in the context of energy-driven collapse models, and as we will see below, most of his analysis also applies to our model.

4.4.1 Maintenance of coherence

First of all, the model satisfies the constraint of predicting the maintenance of coherence when this is observed. Since the energy uncertainty of the state of a microscopic particle is very small in general, its collapse will be too slow

[44]There might exist a subtle connection here. It seems that the energy-conserved wavefunction collapse in discrete time requires a finite event horizon to ensure the energy eigenvalues of any system are discrete. On the other hand, it seems that discrete spacetime permits the existence of dark energy as quantum fluctuations of spacetime to lead to acceleration and finite event horizon (Gao 2005). Besides, the existence of a cosmological constant or extrotical matter also leads to the existence of a finite event horizon.

to have any detectable effect in present experiments on these particles. For example, the energy uncertainty of a photon emitted from an atom is in the order of $10^{-6}eV$, and the corresponding collapse time is $10^{25}s$ according to Eq. (59) of our collapse model, which is much longer than the age of the universe, $10^{17}s$. This means that the collapse states (i.e. energy eigenstates) are never reached for a quantum system with small energy uncertainty even during a time interval as long as the age of the universe. As another example, consider the SQUID experiment of Friedman et al (2000), where the coherent superpositions of macroscopic states consisting of oppositely circulating supercurrents are observed. In the experiment, each circulating current corresponds to the collective motion of about 10^9 Cooper pairs, and the energy uncertainty is about $8.6 \times 10^{-6}eV$. Eq. (59) predicts a collapse time of $10^{23}s$, and thus maintenance of coherence is expected despite the macroscopic structure of the state[45]. For more examples see Adler (2002).

4.4.2 Rapid localization in measurement situations

In the following, we will investigate whether the discrete model of energy-conserved wavefunction collapse can account for the emergence of definite measurement results. Let's first see a simple position measurement experiment. Consider an initial state describing a particle in a superposition of two locations (e.g. a superposition of two gaussian wavepackets separated by a certain distance). After the measurement interaction, the position measuring device evolves to a superposition of two macroscopically distinguishable states:

$$(c_1\psi_1 + c_2\psi_2)\varphi_0 \rightarrow c_1\psi_1\varphi_1 + c_2\psi_2\varphi_2 \tag{64}$$

where ψ_1, ψ_2 are the states of the particle in different locations, φ_0 is the initial state of the position measuring device, and φ_1, φ_2 are the different outcome states of the device. For an ideal measurement, the two particle/device states $\psi_1\varphi_1$ and $\psi_2\varphi_2$ have precisely the same energy spectrum. Then it appears that this superposition will not collapse according to the energy-conserved collapse model.

However, this is not the case. The key is to see that the two states of the particle in the superposition are detected in different parts of the measuring device, and they interact with the different atoms or molecules in these parts.

[45] A potentially more promising case is provided by certain long-lived nuclear isomers, which have large energy gaps from their ground states (see Adler 2002 and references therein). For example, the metastable isomer of ^{180}Ta, the only nuclear isomer to exist naturally on earth, has a half-life of more than 10^{15} years and an energy gap of $75keV$ from the ground state. According to Eq. (59), a coherent superposition of the ground state and metastable isomer of ^{180}Ta will spontaneously collapse to either the isomeric state or the ground state, with a collapse time of order 20 minutes. It will be a promising way to test our collaspe model by examining the maintenance of coherence of such a superposition.

Thus we should rewrite the device states explicitly as $\varphi_0 = \chi_A(0)\chi_B(0)$, $\varphi_1 = \chi_A(1)\chi_B(0)$, and $\varphi_2 = \chi_A(0)\chi_B(1)$, where $\chi_A(0)$ and $\chi_B(0)$ denote the initial states of the device in the parts A and B, respectively, and $\chi_A(1)$ and $\chi_B(1)$ denote the outcome states of the device in the parts A and B, respectively. Then we have

$$(c_1\psi_1 + c_2\psi_2)\chi_A(0)\chi_B(0) \to c_1\psi_1\chi_A(1)\chi_B(0) + c_2\psi_2\chi_A(0)\chi_B(1) \qquad (65)$$

This reformulation clearly shows that there exists energy difference between the sub-systems in the different outcome states of the device. Since there is always some kind of measurement amplification from the microscopic state to the macroscopic outcome in the measurement process, there will be a large energy difference between the states $\chi_A(0)$, $\chi_B(0)$ and $\chi_A(1)$, $\chi_B(1)$. As a result, the total energy difference $\Delta E = |\Delta E_A| + |\Delta E_B|$ is also very large, and it will result in the rapid collapse of the above superposition into one of its branches according to the energy-conserved collapse model[46].

Let's see a more realistic example, a photon being detected via photoelectric effect (e.g. by a single-photon avalanche diode). In the beginning of the detection, the spreading spatial wave function of the photon is entangled with the states of a large number of surface atoms of the detector. In each local branch of the entangled state, the total energy of the photon is wholly absorbed by the electron in the local atom interacting with the photon. This is clearly indicated by the term $\delta(E_f - E_i - \hbar\omega)$ in the transition rate of photoelectric effect. The state of the ejecting electron is a (spherical) wavepacket moving outward from the local atom, which average direction and momentum distribution are determined by the momentum and polarization of the photon. The small energy uncertainty of the photon will also be transfered to the ejecting electron[47].

This microscopic effect of ejecting electron is then amplified (e.g. by an avalanche process of atoms) to form a macroscopic signal such as the shift of the pointer of a measuring device. During the amplification process, the energy difference is constantly increasing between the branch in which the photon is absorbed and the branch in which the photon is not absorbed near each atom interacting with the photon. This large energy difference will soon lead to the collapse of the whole superposition into one of the

[46]Since the uncertainty of the total energy of the whole entangled system is still zero, the energy-driven collapse models will predict that no wavefunction collapse happens and no determinate measurement result appears for the above measurement process, which contradicts experimental observations (Pearle 2004).

[47]In more general measurement situations, the measured particle (e.g. electron) is not annihilated by the detector. However, in each local branch of the entangled state of the whole system, the particle also interacts with a single atom of the detector by an ionizing process, and its total energy is also wholly transfered to the atom and the ejecting electrons.

local branches, and thus the photon is only detected locally. Take the single photon detector - avalanche photodiode as a typical example. Its energy consumption is sharply peaked in a very short measuring interval. One type of avalanche photodiode operates at 10^5 cps and has a mean power dissipation of 4mW (Gao 2006). This corresponds to an energy consumption of about $2.5 \times 10^{11} eV$ per measuring interval $10^{-5} s$. By using the collapse time formula Eq. (59), where the energy uncertainty is $\Delta E \approx 2.5 \times 10^{11} eV$, we find the collapse time is $\tau_c \approx 1.25 \times 10^{-10} s$. This collapse time is much smaller than the measuring interval.

One important point needs to be stressed here. Although a measured particle is detected locally in a detector (e.g. the spatial size of its collapse state is in the order of the size of an atom), its wave funtion does not necessarily undergo the position collapse assumed in an ideal position measurement by standard quantum mechanics, and especially, energy can be conserved during the localization process according to our model. The reason can be summarized as follows. The wave funtion of the measured particle is usually a spherical wave (e.g. a spherically symmetric wave function) in three-dimensional space. Its momentum is along the radial direction, but the local and random measurement result distributes along the sphere, perpendicular to the radial direction. During the detection, the measured particle interacts with a single atom of the detector by an ionizing process in each local branch of the entangled state of the whole system including the particle and the atoms in the detector. The particle is usually absorbed by the atom or bound in the atom, and its energy is wholly transfered to the new-formed atom and the ejecting electrons during the ionizing process in each branch. Then the amplification process such as an avalanche process of atoms introduces very large energy difference between the detected branch and the empty branch, and as a result, the whole superposition will soon collapse into one of its local branches in a random way according to the energy-conserved collapse model. After the collapse, the state of the measured particle is localized in the spatial region of one atom. Moreover, since each local branch of the entangled state of the particle and the detector has the same energy spectrum, the collapse process also conserves energy at the individual level.

4.4.3 Emergence of the classical world

Now let's see whether the discrete model of energy-conserved wavefunction collapse is consistent with the our macroscopic experience. It seems that there is an apparent inconsistency here. According to the model, when there is a superposition of a macroscopic object in an identical physical state (an approximate energy eigenstate) at two different, widely separated locations, the superposition does not collapse. The reason is that there is no energy difference between the two branches of the superposition. However, the ex-

istence of such superpositions is obviously inconsistent with our macroscopic experience; the macroscopic objects are localized. This common objection has been basically answered by Adler (2002). The crux of the matter lies in the influences of environment. The collisions and especially the accretions of environmental particles will quickly increase the energy uncertainty of the entangled state of the whole system including the object and environmental particles, and thus the initial superposition will soon collapse to one of the localized branches according to our model. Accordingly, the macroscopic objects can always be localized due to the environmental influences. Note that the energy uncertainty here denotes the sum of the absolute energy uncertainty of each sub-system in the entangled state as defined in our model[48].

As a typical example, we consider a dust particle of radius $a \approx 10^{-5}cm$ and mass $m \approx 10^{-7}g$. It is well known that localized states of macroscopic objects spread very slowly under the free Schrödinger evolution. For instance, for a Gaussian wave packet with initial (mean square) width Δ, the wave packet will spread so that the width doubles in a time $t = 2m\Delta^2/\hbar$. This means that the double time is almost infinite for a macroscopic object. If the dust particle had no interactions with environment and its initial state is a Gaussian wave packet with width $\Delta \approx 10^{-5}cm$, the doubling time would be about the age of the universe. However, if the dust particle is in interaction with environment, the situation turns out to be very different. Although the different components that couple to the environment will be individually incredibly localised, collectively they can have a spread that is many orders of magnitude larger. In other words, the state of the dust particle and the environment could be a superposition of zillions of very well localised terms, each with slightly different positions, and which are collectively spread over a macroscopic distance (Bacciagaluppi 2008). According to Joos and Zeh (1985), the spread in an environment full of thermal radiation only is proportional to mass times the cube of time for large times, namely $(\Delta x)^2 \approx \Lambda m\tau^3$, where Λ is the localization rate depending on the environment, defined by the evolution equation of density matrix $\rho_t(x, x') = \rho_0(x, x')e^{-\Lambda t(x-x')^2}$. For example, if the above dust particle interacts with thermal radiation at $T = 300K$, the localization rate is $\Lambda = 10^{12}$, and the overall spread of its state is of the order of $10m$ after a second (Joos and Zeh 1985). If the dust particle interacts with air molecules, e.g. floating in the air, the spread of its state will be much faster.

Let's see whether the energy-conserved collapse in our model can prevent the above spreading of the wave packet. Suppose the dust particle is in a superposition of two identical localized states that are separated by $10^{-5}cm$

[48]The uncertainty of the total energy of the whole system is still very small even if the influences of environment are counted. Thus no observable collapse happens for the above situation according to the energy-driven collapse models (Pearle 2004).

in space. The particle floats in the air, and its average velocity is about zero. At standard temperature and pressure, one nitrogen molecule accretes in the dust particle, which area is $10^{-10}cm^2$, during a time interval of $10^{-14}s$ in average (Adler 2002). Since the mass of the dust particle is much larger than the mass of a nitrogen molecule, the velocity change of the particle is negligible when compared with the velocity change of the nitrogen molecules during the process of accretion. Then the kinetic energy difference between an accreted molecule and a freely moving molecule is about $\Delta E = \frac{3}{2}kT \approx 10^{-2}eV$. When one nitrogen molecule accretes in one localized branch of the dust particle (the molecule is freely moving in the other localized branch), it will increase the energy uncertainty of the total entangled state by $\Delta E \approx 10^{-2}eV$. Then after a time interval of $10^{-4}s$, the number of accreted nitrogen molecules is about 10^{10}, and the total energy uncertainty is about 10^8eV. According to Eq. (59) in our collapse model, the correponding collapse time is about $10^{-4}s$. Since the two localized states in the superposition have the same energy spectra, the collapse also conserves energy.

In the energy-conserved collapse model, the collapse states are energy eigenstates, and in particular, they are nonlocal momentum eigenstates for free quantum systems. Thus it is indeed counterintuitive that the energy-conserved collapse can make the states of macroscopic objects local. As shown above, this is due to the constant influences of environmental particles. When the spreading of the state of a macroscopic object becomes larger, its interaction with environmental particles will introduce larger energy difference between its different local branches, and this will then collapse the spreading state again into a more localized state. As a result, the states of macroscopic objects in an environment will never reach the collapse states, namely momentum eigenstates, though they do continuously undergo the energy-conserved collapse. To sum up, there are two opposite processes for a macroscopic object constantly interacting with environmental particles. One is the spreading process due to the linear Schrödinger evolution, and the other is the localization process due to the energy-conserved collapse evolution. The interactions with environmental particles not only make the spreading more rapidly but also make the localization more frequently. In the end these two processes will reach an approximate equilibrium. The state of a macroscopic object will be a wave packet narrow in both position and momentum, and this narrow wave packet will follow approximately Newtonian trajectories (if the external potential is uniform enough along the width of the packet) by Ehrenfest's theorem (See Bacciagaluppi 2008 for a similar analysis in the context of decoherence)[49]. In some sense, the

[49]When assuming the energy uncertainty of an object is in the same order of its thermal energy fluctuation, we can estimate the rough size of its wavepacket. For instance, for a dust particle of mass $m = 10^{-7}g$, its root mean square energy fluctuation is about 10^3eV at room temperature $T = 300K$ (Adler 2002), and thus the width of its wavepacket is about $10^{-10}m$.

emergence of the classical world around us is "conspired" by environmental particles according to the energy-conserved collapse model.

In the ultimate, the energy-conserved collapse model should be able to account for our definite conscious experience. According to the existing neuroscience knowledge, the appearance of a (definite) conscious perception in human brains involves a large number of neurons changing their states from resting state (resting potential) to firing state (action potential). In each neuron, the main difference of these two states lies in the motion of 10^6 Na^+s passing through the neuron membrane. Since the membrane potential is in the order of $10^{-2}V$, the energy difference between firing state and resting state is $\Delta E \approx 10^4 eV$. According to the energy-conserved collapse model, the collapse time of a quantum superposition of these two states of a neuron is

$$\tau_c \approx \frac{\hbar E_P}{(\Delta E)^2} \approx (\frac{2.8MeV}{0.01MeV})^2 \approx 10^5 s, \tag{66}$$

where the Planck energy $E_P \approx 10^{19} GeV$. When considering the number of neurons that can form a definite conscious perception is usually in the order of 10^7, the collapse time of the quantum superposition of two different conscious perceptions will be

$$\tau_c \approx (\frac{2.8MeV}{100GeV})^2 \approx 10^{-9} s. \tag{67}$$

Since the normal conscious time of a human being is in the order of several hundred milliseconds, the collapse time is much shorter than the normal conscious time. Therefore, our conscious perceptions are always definite according to the energy-conserved collapse model.

4.5 Critical comments on other dynamical collapse models

In this subsection, we will give a critical analysis of other dynamical collapse models. These models can be sorted into two categories. The first one may be called spontaneous collapse models, in which the dynamical collapse of the wave function is assumed to happen even for an isolated system. They include the gravity-induced wavefunction collapse model (Diósi 1989; Penrose 1996), the GRW model (Ghirardi, Rimini and Weber 1986)[50] etc. The second category may be called interaction-induced collapse models, which

[50]The GRW model was originally referred to as QMSL (Quantum Mechanics with Spontaneous Localizations). In this model, it is assumed that each elementary constituent of any physical system is subjected, at random times, to random and spontaneous localization processes (or hittings) around appropriate positions. The random hittings happen much less frequently for a microscopic system, e.g. an electron undergoes a hitting, on average, every hundred million years. If these hittings are assumed to be brought by an external system, then the GRW model should be regarded not as a spontaneous collapse model but as an interaction-induced collapse model.

assume that the dynamical collapse of the wave function of a given system results from its particular interaction with a noise field. One typical example is the CSL model (Pearle 1989; Ghirardi, Pearle and Rimini 1990)[51]. In the following, we will mainly analyze Penrose's gravity-induced wavefunction collapse model and the CSL model, which are generally regarded as two of the most promising models of wavefunction collapse.

4.5.1 Penrose's gravity-induced wavefunction collapse model

It seems very natural to guess that the collapse of the wave function is induced by gravity. The reasons include: (1) gravity is the only universal force being present in all physical interactions; (2) gravitational effects grow with the size of the objects concerned, and it is in the context of macroscopic objects that linear superpositions may be violated. Though the gravity-induced collapse conjecture may be traced back to Feynman (1995), it is Penrose (1996) who proposed a concrete gravity-induced collapse argument.

Penrose's argument is based on a profound and fundamental conflict between the general covariance principle of general relativity and the superposition principle of quantum mechanics. The conflict can be clearly seen by considering the superposition state of a static mass distribution in two different locations, say position A and position B. On the one hand, according to quantum mechanics, the valid definition of such a superposition requires the existence of a definite space-time background, in which position A and position B can be distinguished. On the other hand, according to general relativity, the space-time geometry, including the distinguishability of position A and position B, cannot be predetermined, and must be dynamically determined by the position superposition state. Since the different position states in the superposition determine different space-time geometries, the space-time geometry determined by the whole superposition state is indefinite, and as a result, the superposition and its evolution cannot be consistently defined. In particular, the definition of the time-translation operator for the superposed space-time geometries involves an inherent ill-definedness, and this leads to an essential uncertainty in the energy of the superposed state. Then by analogy Penrose argued that this superposition, like an unstable particle in usual quantum mechanics, is also unstable, and it will decay or collapse into one of the two states in the superposition after a finite lifetime. Furthermore, Penrose suggested that the essential energy uncertainty in the Newtonian limit is proportional to the gravitational self-energy E_Δ of the difference between the two mass distributions, and the collapse time, analogous to the half-life of an unstable particle, is

$$T \approx \hbar / E_\Delta. \tag{68}$$

[51] If the involved noise field in the CSL model is not taken as real, then the model should be regarded as a spontaneous collapse model.

This criterion is very close to that put forward by Diósi (1989) earlier, and it is usually called the Diósi-Penrose criterion. Later, Penrose (1998) further suggested that the collapse states are the stationary solutions of the Schrödinger-Newton equation.

Let's now analyze Penrose's argument. The crux of the matter is whether the conflict between quantum mechanics and general relativity requires that a quantum superposition of two space-time geometries must collapse after a finite time. We will argue in the following that the answer is probably negative. First of all, although it is widely acknowledged that there exists a fundamental conflict between the general covariance principle of general relativity and the superposition principle of quantum mechanics, it is still a controversial issue what the exact nature of the conflict is and especially how to solve it. For example, it is also possible that the conflict may be solved by reformulating quantum mechanics in a way that does not rely on a definite spacetime background (see, e.g. Rovelli 2011).

Next, Penrose's argument seems too weak to establish a necessary connection between the conflict and wavefunction collapse. Even though there is an essential uncertainty in the energy of the superposition of different space-time geometries, this kind of energy uncertainty is different in nature from the energy uncertainty of unstable particles or unstable states in usual quantum mechanics (Gao 2010). The former results from the ill-definedness of the time-translation operator for the superposed space-time geometries (though its nature seems still unclear), while the latter exists in a definite spacetime background, and there is a well-defined time-translation operator for the unstable states. Moreover, the decay of these unstable states is a natural result of the linear Schrödinger evolution, and the process is not random but deterministic. By contrast, the hypothetical spontaneous decay or collapse of the superposed space-time geometries is nonlinear and random. In addition, the decay of an unstable state (e.g. excited state of an atom) is actually not spontaneous but caused by the background field constantly interacting with it. In some extreme situations, the state may not decay at all when being in a very special background field with bandgap (Yablonovitch 1987). In short, there exists no convincing analogy between a superposition of different space-time geometries and an unstable state in usual quantum mechanics. Accordingly, one cannot argue for the decay or collapse of the superposition of different space-time geometries by this analogy. Although an unstable state in quantum mechanics may decay after a very short time, this does not imply that a superposition of different space-time geometries should also decay, let alone sometimes an unstable state does not decay at all under a special circumstance. To sum up, Penrose's argument by analogy only has a very limited force, and especially, it is not strong enough to establish a necessary connection between the conflict between quantum mechanics and general relativity and wavefunction collapse.

Thirdly, it can be further argued that the conflict does not necessarily

lead to wavefunction collapse. The key is to realize that the conflict also needs to be solved before the wavefunction collapse finishes, and when the conflict has been solved, the wavefunction collapse will lose its basis relating to the conflict. As argued by Penrose (1996), the quantum superposition of different space-time geometries and its evolution are both ill-defined due to the fundamental conflict between the general covariance principle of general relativity and the superposition principle of quantum mechanics. The ill-definedness seems to require that the superposition must collapse into one of the definite space-time geometries, which has no problem of such ill-definedness. However, the wavefunction collapse seems too late to save the superposition from the "suffering" of the ill-definedness during the collapse. In the final analysis, the conflict or the problem of ill-definedness needs to be solved *before* defining a quantum superposition of different space-time geometries and its evolution. In particular, the possible collapse evolution of the superposition also needs to be consistently defined, which again indicates that the wavefunction collapse does not solve the problem of ill-definedness. On the other hand, once the problem of ill-definedness is solved and a consistent description obtained (however this is still an unsovled issue in quantum gravity), the wavefunction collapse will completely lose its connection with the problem[52]. Therefore, contrary to Penrose's expectation, it seems that the conflict between quantum mechanics and general relativity does not necessarily entail the existence of wavefunction collapse.

Even though Penrose's gravity-induced collapse argument is debatable, the wavefunction collapse may still exist due to other reasons, and thus Penrose's concrete suggestions for the collapse time formula and collapse states also need to be further examined as some aspects of a phenomenological model. First of all, let's analyze Penrose's collapse time formula Eq. (68), according to which the collapse time of a superposition of two mass distributions is inversely proportional to the gravitational self-energy of the difference between the two mass distributions. As we have argued above, a convincing analogy between such a superposition and an unstable state in quantum mechanics does not exist, and gravity does not necessarily induce wavefunction collapse either. Thus this collapse time formula, which is based on a similar application of Heisenberg's uncertainty principle to unsta-

[52]Note that if the problem of ill-definedness cannot be solved in principle for the superpositions of very different space-time geometries, then the wavefunction collapse may be relevant here. Concretely speaking, if the superpositions of very different space-time geometries cannot be consistently defined in nature, then it is very likely that these superpositions cannot exist, which means that they must have collapsed into one of the definite space-time geometries before formed from the superpositions of minutely different space-time geometries. In this case, the large difference of the space-time geometries in the superposition will set a upper limit for wavefunction collapse. Though the limit may be loose, it does imply the existence of wavefunction collapse. However, this possibility might be very small, as it seems that there is always some kind of approximate sense in which two different spacetimes can be pointwise identified.

ble states, will lose its original physical basis. In particuar, the appearance of the gravitational self-energy term in the formula is in want of a reasonable explanation. In fact, it has already been shown that this gravitational self-energy term does not represent the ill-definedness of time-translation operator (or the fuzziness of the identification between two spacetimes) in the strictly Newtonian regime (Christian 2001). In this regime, the time-translation operator can be well defined, but the gravitational self-energy term is not zero. In addition, as Diósi (2007) pointed out, the microscopic formulation of the collapse time formula is unclear and still has some problems (e.g. the cut-off difficulty).

Next, let's examine Penrose's suggestion for the collapse states. According to Penrose (1998), the collapse states are the stationary solutions of the Schrödinger-Newton equation, namely Eq. (5) given in Section 2. The equation describes the gravitational self-interaction of a single quantum system, in which the mass density $m|\psi(x,t)|^2$ is the source of the classical gravitational potential. As we have argued in Section 2, although a quantum system has mass density that is measurable by protective measurement, the density is not real but effective, and it is formed by the ergodic motion of a localized particle with the total mass of the system. Therefore, there does not exist a gravitational self-interaction of the mass density. This conclusion can also be reached by another somewhat different argument. Since charge always accompanies mass for a charged particle such as an electron[53], the existence of the gravitational self-interaction, though which is too weak to be excluded by present experiments, may further entail the existence of a remarkable electrostatic self-interaction of the particle, which already contradicts experiments as we have shown in Section 2. This analysis poses a serious objection to the Schrödinger-Newton equation and Penrose's suggestion for the collapse states[54].

Lastly, we briefly discuss another two problems of Penrose's collapse scheme. The first one is the origin of the randomness of collapse results. Penrose did not consider this issue in his collapse scheme. If the collapse is indeed spontaneous as implied by his gravity-induced collapse argument, then the randomness cannot result from any external influences such as an external noise field, and it can only come from the studied quantum system and its wave function. The second problem is energy non-conservation. Although Penrose did not give a concrete model of wavefunction collapse, his

[53]However, the concomitance of mass and charge in space for a charged particle does not necessarily require that they must satisfy the same law of interaction. For example, the fact that electromagnetic fields are quantized in nature does not necessarily imply that gravitational fields must be also quantized.

[54]Since the Schrödinger-Newton equation is the non-relativistic realization of the typical model of semiclassical gravity, in which the source term in the classical Einstein equation is taken as the expectation of the energy momentum operator in the quantum state, our analysis also presents a serious objection to the approach of semiclassical gravity.

collapse scheme requires the collapse of superpositions of different positions, while this kind of space collapse inevitably violates energy conservation[55]. Since the gravitational energy of a quantum system is much smaller than the energy of the system, Penrose's collapse scheme still violates energy conservation even if the gravitational field is counted. As we have noted earlier, for an isolated system only the collapse states are energy eigenstates can energy conserve (at the ensemble level) during the collapse. If the principle of conservation of energy is indeed universal as widely thought, then the spontaneous collapse models that violate enery conservation will have been excluded. By contrast, although the interaction-induced collapse models such as the CSL model also violate energy conservation in their present formulations, there is still hope that when counting the energy of external noise field the total energy may be conserved in these models (Pearle 2000; Bassi, Ippoliti and Vacchini 2005). Let's turn to the CSL model now.

4.5.2 The CSL model

In the CSL model, the collapse of the wave function of a quantum system is assumed to be caused by its interaction with a classical scalar field, $w(x,t)$. The collapse states are the eigenstates of the smeared mass density operator, and the mechanism leading to the suppression of the superpositions of macroscopically different states is fundamentally governed by the integral of the squared differences of the mass densities associated to the superposed states. It may be expected that the introduction of the noise field can help to solve the problems plagued by the spontaneous collapse models, e.g. the problems of energy non-conservation and the origin of randomness etc. However, one must first answer what field the noise field is and especially why it can collapse the wave functions of all quantum systems. The validity of the CSL model strongly depends on the existence of this hypothetical noise field. In this subsection, we will mainly analyze this important legitimization problem of the CSL model[56].

Whatever the nature of the noise field $w(x,t)$ is, it cannot be quantum in the usual sense since its coupling to a quantum system is not a standard coupling between two quantum systems. The coupling is anti-Hermitian (Bassi 2007), and the equation of the resulting dynamical collapse is not the

[55]Diósi (2007) explicitly pointed out that the von-Neumann-Newton equation, which may be regarded as one realization of Penrose's collapse scheme, obviously violates conservation of energy.

[56]Related to this legitimization problem is that the two parameters which specify the model are ad hoc (Pearle 2007). These two parameters, which were originally introduced by Ghirardi, Rimini and Weber (1986), are a distance scale, $a \approx 10^5 cm$, characterising the distance beyond which the collapse becomes effective, and a time scale, $\lambda^{-1} \approx 10^{16}$sec, giving the rate of collapse for a microscopic system. If wavefunction collapse is a fundamental physical process related to other fundamental processes, the parameters should be able to be written in terms of other physical constants.

standard Schrödinger equation with a stochastic potential either. According to our current understandings, the gravitational field is the only universal field that might be not quantized, though this possibility seems extremely small in the view of most researchers. Therefore, it seems natural to identify this noise field with the gravitational field. In fact, it has been argued that in the CSL model the w-field energy density must have a gravitational interaction with ordinary matter (Pearle and Squires 1996; Pearle 2009). The argument of Pearle and Squires (1996) can be summarized as follows[57].

There are two equations which characterize the CSL model. The first equation is a modified Schrödinger quation, which expresses the influence of an arbitrary field $w(x,t)$ on the quantum system. The second equation is a probability rule which gives the probability that nature actually chooses a particular $w(x,t)$. This probability rule can also be interpreted as expressing the influence of the quantum system on the field. As a result, $w(x,t)$ can be written as follows:

$$w(x,t) = w_0(x,t) + < A(x,t) >, \tag{69}$$

where $A(x,t)$ is the mass density operator smeared over the GRW scale a, $< A(x,t) >$ is its quantum expectation value, and $w_0(x,t)$ is a gaussian randomly fluctuating field with zero drift, temporally white noise in character and with a particular spatial correlation function. Then the scalar field $w(x,t)$ that causes collapse can be interpreted as the gravitational curvature scalar with two sources, the expectation value of the smeared mass density operator and an independent white noise fluctuating source. This indicates that the CSL model is based on the semi-classical gravity, and the smeared mass density is the source of the gravitational potential. Note that the reality of the field $w(x,t)$ requires that the smeared mass density of a quantum system is real.

According to our previous analysis, however, a quantum system does not have a real mass density distribution in space, no matter it is smeared or not. Moreover, although the approach of semi-classical gravity may be consistent in the context of dynamical collapse models (Pearle and Squires 1996; Ghirardi 2008), it may have been excluded as implied by the analysis. Besides, as we have pointed out in Section 2, protective measurement shows that a quantum system has an effective mass density proportional to the modulus square of its wave function. Thus the assumed existence of the *smeared* mass density in the CSL model, even if it is effective, also contradicts protective measurement. Note that it is crucial that the mass density be smeared over the GRW scale a in the CSL model; without such a smearing the energy excitation of particles undergoing collapse would be beyond experimental constraints (Pearle and Squires 1996). In conclusion,

[57] Pearle (2009) further argued that compatibility with general relativity requires a gravitational force exerted upon matter by the w-field.

it seems that the noise field introduced in the CSL model cannot have a gravitational origin required by the model, and this may raise strong doubts about the reality of the field.

On the other hand, even though the semi-classical gravity is viable and the noise field in the CSL model can be the gravitational field, one still need to answer why the gravitational field has the very ability to collapse the wave functions of all quantum systems as required by the model. It is worth noting that the randomly fluctuating field in the model, $w_0(x,t)$, is not the gravitational field of the studied quantum system but the background gravitational field. Thus Penrose's gravity-induced wavefunction collapse argument, even if valid, does not apply to the CSL model, which is essentially an interaction-induced model of wavefunction collapse. The fluctuations of the background gravitational field can readily lead to the decoherence of the wave function of a quantum system, but it seems that they have no ability to cause the collapse of the wave function.

Lastly, let's briefly discuss another two problems of the CSL model. The first one is the well-known problem of energy non-conservation. The collapse in the model narrows the wave function in position space, thereby producing an increase of energy[58]. A possible solution is that the conservation laws may be satisfied when the contributions of the noise field $w(x,t)$ to the conserved quantities are taken into account. It has been shown that the total mean energy can be conserved (Pearle 2004), and the energy increase can also be made finite when further revising the coupling between the noise field and the studied quantum system (Bassi, Ippoliti and Vacchini 2005). But a complete solution has not been found yet, and it is still unknown whether such a solution indeed exists. The second problem is to make a relativistic quantum field theory which describes collapse (Pearle 2009). Notwithstanding a good deal of effort, a satisfactory theory has not been obtained at present (see Bedingham 2011 for a recent attempt). The main difficulty is that the hypothetical interaction responsible for collapse will produce too many particles out of the vacuum, amounting to infinite energy per sec per volume, in the relativistic extension of these interaction-induced collapse models. Note that the spontaneous collapse models without collapse interaction (e.g. the energy-conserved collapse model) is lack of this difficulty.

5 Conclusions

In this paper, we argue that the mass and charge density of a quantum system, which is measurable by protective measurement and proportional to the modulus square of its wave function, is formed by the random and discontinuous ergodic motion of a localized particle with the total mass and

[58]Note that with appropriate choice for the parameters in the CSL model, such a violation of energy conservation is very tiny and hardly detectable by present day technology.

charge of the system. This result suggests a new interpretation of the wave function, according to which the wave function on configuration space is a description of random discontinuous motion of particles in the real three-dimensional space, and the modulus square of the wave function gives the probability density of the particles being in certain locations in space. We show that the suggested interpretation of the wave function is consistent with the derivation of the free Schrödinger equation based on spacetime translation invariance and relativistic invariance, and more importantly, it also has some implications for the solution to the measurement problem; the de Broglie-Bohm theory and the many-worlds interpretation are disfavored, and the collapse of the wave function is probably a real and dynamical process. It is further argued that the random discontinuous motion of particles may be the physical origin of the wavefunction collapse, and a discrete model of energy-conserved wavefunction collapse is also proposed and shown consistent with existing experiments and our macroscopic experience.

The new analysis of the meaning of the wave function and its evolution may have some possible implications for the unification of quantum mechanics and special and general relativity. First of all, a consistent relativistic description of random discontinuous motion seems to require the existence of a preferred Lorentz frame, and the collapse dynamics also provides a way to detect the frame according to our collapse model. This suggests a possible solution to the problem of incompatibility between quantum nonlocality and the principle of relativity. Next, the existence of a preferred Lorentz frame may also help to settle the controversial issue concerning the physical explanation of quantum field. Since the preferred Lorentz frame can restore the absoluteness of simultaneity, the existing objections to the particle interpretation of quantum field may be avoided. For example, it permits the existence of local number operators and a unique total number operator even for interacting quantum fields (Bain 2011). As a result, the quantum field theory may still be regarded as a theory describing the relativistic motion of particles, including the creation and annihilation of particles as a special kind of motion. Lastly, the energy-conserved wavefunction collapse, if indeed exists, might help to fulfill the unification of quantum mechanics and general relativity. In particular, it will prohibit the existence of superpositions of very different spacetime geometries and may also prevent the formation of singularities, and thus it seems possible that the gravitational field may be not quantized in the standard way. Moreover, different from the semi-classical approach of quantum gravity, the discrete collapse dynamics might further provide a consistent framework for a fundamental theory of quantum gravity. A detailed analysis of these possible implications will be given in a future paper.

Acknowledgments

This work was supported by the Postgraduate Scholarship in Quantum Foundations provided by the Unit for History and Philosophy of Science and Centre for Time (SOPHI) of the University of Sydney. I am very grateful to Dean Rickles and Huw Price for their continuous support, stimulating suggestions and healthy criticism which helped me greatly during my research.

References

[1] Adler, S. L. (2002). Environmental influence on the measurement process in stochastic reduction models, J. Phys. A: Math. Gen. 35, 841-858.

[2] Aharonov, Y., Anandan, J. and Vaidman, L. (1993). Meaning of the wave function, Phys. Rev. A 47, 4616.

[3] Aharonov, Y., Anandan, J. and Vaidman, L. (1996). The meaning of protective measurements, Found. Phys. 26, 117.

[4] Aharonov, Y. and Vaidman, L. (1993). Measurement of the Schrödinger wave of a single particle, Phys. Lett. A 178, 38.

[5] Allori, V., Goldstein, S., Tumulka, R., and Zanghì, N. (2008), On the Common Structure of Bohmian Mechanics and the Ghirardi-Rimini-Weber Theory, British Journal for the Philosophy of Science 59 (3), 353-389.

[6] Anandan, J. (1993). Protective Measurement and Quantum Reality. Found. Phys. Lett., 6, 503-532.

[7] Bacciagaluppi, G. (2008). The role of decoherence in quantum mechanics. The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.). http://plato.stanford.edu/archives/fall2008/entries/qm-decoherence.

[8] Bain, J. (2011). Quantum field theories in classical spacetimes and particles. Studies in History and Philosophy of Modern Physics 42, 98-106.

[9] Bassi, A., Ippoliti, E., and Vacchini, B. (2005). On the energy increase in space-collapse models. J. Phys. A : Math. Gen. 38, 8017.

[10] Bassi, A. Dynamical reduction models: present status and future developments. J. Phys.: Conf. Series 67, 012013 (2007).

[11] Bedingham, D. J. (2011). Relativistic state reduction dynamics. Found. Phys. 41, 686-704.

[12] Bell, J. (1990) Against measurement, in A. I. Miller (ed.), Sixty-Two Years of Uncertainty: Historical Philosophical and Physics Enquiries into the Foundations of Quantum Mechanics. Berlin: Springer, 17-33.

[13] Belot, G. (2011). Quantum states for primitive ontologists: A case study. http://philsci-archive.pitt.edu/8468/. Forthcoming in European Journal for Philosophy of Science.

[14] Bohm, D. (1952). A suggested interpretation of quantum theory in terms of "hidden" variables, I and II. Phys. Rev. 85, 166-193.

[15] Bohm D. and Hiley, B.J. (1993) The Undivided Universe: An Ontological Interpretation of Quantum Theory. London: Routledge.

[16] Brown, H. R., Dewdney, C. and Horton, G. (1995) Bohm particles and their detection in the light of neutron interferometry. Found. Phys. 25, 329.

[17] Christian, J. (2001). Why the quantum must yield to gravity. In: Physics Meets Philosophy at the Planck Scale, C. Callender and N. Huggett (eds.). Cambridge: Cambridge University Press. pp.305.

[18] Czachor, M. (1996). Nonlinear Schrödinger equation and two-level atoms. Phys. Rev. A 53, 1310.

[19] de Broglie, L. (1928). in: Electrons et photons: Rapports et discussions du cinquime Conseil de Physique Solvay, eds. J. Bordet. Paris: Gauthier-Villars. pp.105. English translation: G. Bacciagaluppi and A. Valentini (2009), Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference. Cambridge: Cambridge University Press.

[20] Diósi, L. (1984). Gravitation and the quantum-mechanical localization of macro-objects. Phys. Lett. A 105, 199-202.

[21] Diósi, L. (1989). Models for universal reduction of macroscopic quantum fluctuations. Phys. Rev. A 40, 1165-1173.

[22] Diósi, L. (2007). Notes on certain Newton gravity mechanisms of wave function localisation and decoherence. J. Phys. A: Math. Gen. 40, 2989-2995.

[23] Duff, M. J., Okun, L. B., and Veneziano, G. (2002). Trialogue on the number of fundamental constants. Journal of High Energy Physics, 0203, 23.

[24] Everett, H. (1957). 'Relative state' formulation of quantum mechanics. Rev. Mod. Phys. 29, 454-462.

[25] Feynman, R. (1995). Feynman Lectures on Gravitation. B. Hatfield (ed.), Reading, Massachusetts: Addison-Wesley.

[26] Friedman J. R., Patel, V., Chen, W., Tolpygo, S. K. and Lukens, J. E. (2000). Nature 406, 43.

[27] Gao, S. (2004). Quantum collapse, consciousness and superluminal communication, Found. Phys. Lett, 17, 167-182.

[28] Gao, S. (2005). A conjecture on the origin of dark energy. Chinese Physics Letters 22, 783.

[29] Gao, S. (2006). A model of wavefunction collapse in discrete space-time. Int. J. Theor. Phys. 45, 1943-1957.

[30] Gao, S. (2010). On Diósi-Penrose criterion of gravity-induced quantum collapse. Int. J. Theor. Phys. 49, 849-853.

[31] Gao, S. (2011). Protective measurement and the meaning of the wave function. http://philsci-archive.pitt.edu/8741/.

[32] Garay, L. J. (1995). Quantum gravity and minimum length. International Journal of Modern Physics A 10, 145.

[33] Ghirardi, G. C. (1997), Quantum Dynamical Reduction and Reality: Replacing Probability Densities with Densities in Real Space, Erkenntnis, 45, 349.

[34] Ghirardi, G. C. (2008). Collapse Theories, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.). http://plato.stanford.edu/archives/fall2008/entries/qm-collapse.

[35] Ghirardi, G. C., Grassi, R. and Benatti, F. (1995). Describing the macroscopic world: Closing the circle within the dynamical reduction program. Found. Phys., 25, 313328.

[36] Ghirardi, G. C., Pearle, P. and Rimini, A. (1990). Markov-processes in Hilbert-space and continuous spontaneous localization of systems of identical particles. Phys. Rev. A 42, 78.

[37] Ghirardi, G.C., Rimini, A., and Weber, T. (1986). Unified dynamics for microscopic and macroscopic systems. Phys. Rev. D 34, 470.

[38] Goldstein, S. (2009). Bohmian Mechanics, The Stanford Encyclopedia of Philosophy (Spring 2009 Edition), Edward N. Zalta (ed.). http://plato.stanford.edu/archives/spr2009/entries/qm-bohm.

[39] Hanson, H. and Thoma, F. (2011) Bohmian Mechanics, https://wiki.physik.uni-muenchen.de/TMP/images/b/b4/BohmianMechnics.pdf. Accessed on: July 21, 2011.

[40] Holland, P. (1993). The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics. Cambridge: Cambridge University Press.

[41] Holman, M. (2006) On arguments for linear quantum dynamics. arXiv: quant-ph/0612209.

[42] Hughston, L. P. (1996). Geometry of stochastic state vector reduction. Proc. Roy. Soc. A 452, 953.

[43] Joos, E. and Zeh, H. D. (1985). The emergence of classical properties through interaction with the environment. Zeitschrift fr Physik B 59, 223-243.

[44] Lewis, P. (2004). Life in configuration space, British Journal for the Philosophy of Science 55, 713729.

[45] Monton, B. (2002). Wave function ontology. Synthese 130, 265-277.

[46] Nelson, E. (1966). Derivation of the Schrödinger equation from Newtonian mechanics. Phys. Rev. 150, 10791085.

[47] Okun, L. B. (2009). Mass versus relativistic and rest masses. Am. J. Phys. 77, 430.

[48] Pearle, P. (1989). Combining stochastic dynamical state-vector reduction with spontaneous localization. Phys. Rev. A 39, 2277.

[49] Pearle, P. (1999). Collapse models. In: Petruccione, F. and Breuer, H. P. (eds.), Open Systems and Measurement in Relativistic Quantum Theory. Springer Verlag, New York.

[50] Pearle, P. (2000). Wavefunction Collapse and Conservation Laws. Found. Phys. 30, 1145-1160.

[51] Pearle, P. (2004). Problems and aspects of energy-driven wavefunction collapse models. Phys. Rev. A 69, 42106.

[52] Pearle, P. (2007). How stands collapse I. J. Phys. A: Math. Theor., 40, 3189-3204.

[53] Pearle, P. (2009). How stands collapse II. in Myrvold, W. C. and Christian, J. eds., Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony. The University of Western Ontario Series in Philosophy of Science, 73(IV), 257-292.

[54] Pearle, P. and Squires, E. (1996). Gravity, energy conservation and parameter values in collapse models. Found. Phys. 26, 291.

[55] Penrose, R. (1996). On gravity's role in quantum state reduction. Gen. Rel. Grav. 28, 581.

[56] Penrose, R. (1998). Quantum computation, entanglement and state reduction. Phil. Trans. R. Soc. Lond. A 356, 1927.

[57] Percival, I. C. (1995). Quantum space-time fluctuations and primary state diffusion. Proc. Roy. Soc. A 451, 503.

[58] Percival, I. C. (1998). Quantum State Diffusion. Cambridge: Cambridge University Press.

[59] Rovelli, C. (2011). "Forget time": Essay written for the FQXi contest on the Nature of Time. Found. Phys. 41, 1475-1490.

[60] Salzman, P. J. (2005). Investigation of the Time Dependent Schrödinger-Newton Equation, Ph.D. Dissertation, University of California at Davis.

[61] Saunders, S., Barrett, J., Kent, A. and Wallace, D. (eds.) (2010) Many Worlds? Everett, Quantum Theory, and Reality. Oxford: Oxford University Press.

[62] Shankar, R. (1994). Principles of Quantum Mechanics, 2nd ed. New York: Plenum.

[63] Sonego, S. and Pin, M. (2005). Deriving relativistic momentum and energy. Eur. J. Phys. 26, 33-45.

[64] Suárez, M. (2004), Quantum selections, propensities and the problem of measurement, British Journal for the Philosophy of Science, 55(2), 219-55.

[65] Tooley, M. (1988). In defence of the existence of states of motion. Philosophical Topics 16, 225-254.

[66] Vaidman, L. (2009). Protective measurements, in Greenberger, D., Hentschel, K., and Weinert, F. (eds.), Compendium of Quantum Physics: Concepts, Experiments, History and Philosophy. Springer-Verlag, Berlin. pp.505-507.

[67] Valentini, A. (1997). On Galilean and Lorentz invariance in pilot-wave dynamics. Phys. Lett. A 228, 215222.

[68] Wallace, D. and Timpson, C. G. (2010). Quantum mechanics on space-time I: Spacetime state realism. British Journal for the Philosophy of Science, 61 (4): 697-727.

[69] Yablonovitch, E. (1987). Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett. 58, 2059.