Protective Measurement and the Meaning of the Wave Function

Shan Gao*

November 27, 2013

Abstract

This article introduces the method of protective measurement and discusses its possible implications for the meaning of the wave function. It is argued that the results of protective measurements as predicted by quantum mechanics imply that the wave function of a quantum system is a representation of the physical state of the system, and a further analysis of the mass and charge distribution of the system, which is measurable by protective measurements, may also help determine what physical state the wave function represents.

1 Introduction

Protective measurement, in the language of standard quantum mechanics, is a method to measure the expectation value of an observable on a single quantum system (Aharonov and Vaidman 1993; Aharonov, Anandan and Vaidman 1993; Aharonov, Anandan and Vaidman 1996; Vaidman 2009). For a conventional measurement such as an impulsive measurement, the state of the measured system is entangled with the state of the measuring device during the measurement. By contrast, during a protective measurement the measured state is protected by an appropriate procedure so that it neither changes nor becomes entangled with the state of the measuring device appreciably. In this way, such protective measurements can measure the expectation values of observables on a single quantum system, even if the system is initially not in an eigenstate of the measured observable, and in particular, the wave function of the system can also be measured as expectation values of certain observables.

In this article, we will introduce the method of protective measurement and discuss its possible implications for the foundations of quantum mechanics, especially for the physical meaning of the wave function. Section 2 presents a strict formulation of protective measurement. Sections 3 and 4 then analyze the physical implications of protective measurement. It is argued that the results of protective measurements as predicted by quantum mechanics imply that the wave function of a quantum system is a representation of the physical state of

*Institute for the History of Natural Sciences, Chinese Academy of Sciences, Beijing 100190, P. R. China. E-mail: gaoshan@ihns.ac.cn.

the system, and a further analysis of the mass and charge distribution of the system, which is measurable by a series of protective measurements, may also help determine what physical state the wave function represents. Conclusions are given in the last section.

2 Protective measurements

For a conventional measurement such as an impulsive measurement, if the measured system, prior to the measurement of a variable A, is not in an eigenstate of A, then its state will be invariably entangled with the state of the device due to the interaction. A protective measurement differs from the conventional measurement in that the measured state is protected from entangling and changing appreciably when the measurement is being made[1]. A universal protection scheme is via the quantum Zeno effect. Let's see how this can be done.

Let $|\psi\rangle$ be the measured state of a single system at a given instant $t = 0$. To protect this state from being changed, we make projective measurements of an observable $P(t)$, for which $|\psi\rangle$ is a nondegenerate eigenstate, a large number of times which are dense in the measurement interval $[0, T]$ (Aharonov, Anandan and Vaidman 1993). For example, $P(t)$ is measured in $[0, T]$ at times $t_n = (n/N)T, n = 1, 2, ..., N$, where N is an arbitrarily large number. At the same time, an observable A is measured in the interval $[0, T]$ by an independent measurement described by the following Hamiltonian:

$$H(t) = H_S + H_D + g(t)PA, \tag{1}$$

where H_S and H_D are the free Hamiltonians of the measured system and the measuring device, respectively, and P is the momentum conjugate to the pointer variable X of the device. The time-dependent coupling strength $g(t)$ is a smooth function normalized to $\int dtg(t) = 1$ in $[0, T]$, and $g(0) = g(T) = 0$. The initial state of the pointer at $t = 0$, $|\phi(0)\rangle$, is supposed to be a Gaussian wave packet of eigenstates of X with width w_0, centered around the eigenvalue x_0.

Then the branch of the state of the combined system after T, in which each projective measurement of $P(t_n)$ results in the state of the measured system being in $|\psi\rangle$, is given by

$$\begin{aligned} |t = T\rangle &= |\psi\rangle \langle\psi|e^{-\frac{i}{\hbar}\frac{T}{N}H(t_N)}... |\psi\rangle \langle\psi|e^{-\frac{i}{\hbar}\frac{T}{N}H(t_2)} |\psi\rangle \langle\psi| \\ &\quad \times e^{-\frac{i}{\hbar}\frac{T}{N}H(t_1)} |\psi\rangle |\phi(0)\rangle \\ &= |\psi\rangle \langle\psi|e^{-\frac{i}{\hbar}\frac{T}{N}g(t_N)PA}... |\psi_1\rangle \langle\psi|e^{-\frac{i}{\hbar}\frac{T}{N}g(t_2)PA} |\psi_1\rangle \langle\psi| \\ &\quad \times e^{-\frac{i}{\hbar}\frac{T}{N}g(t_1)PA} |\psi_1\rangle |\phi(T)\rangle , \end{aligned} \tag{2}$$

where $|\psi_1\rangle$ is the state of the system after it evolves from the state $|\psi\rangle$ under the Hamiltonian H_S for a time interval T/N, and $|\phi(T)\rangle$ is the state of the device when it evolves under the Hamiltonian H_D after T. Here it is assumed that

[1]It might be worth noting that there appeared numerous objections to the validity of protective measurements (see, e.g. Unruh 1994; Rovelli 1994; Ghose and Home 1995; Uffink 1999), and these objections have been answered (Aharonov, Anandan and Vaidman 1996; Dass and Qureshi 1999; Vaidman 2009; Gao 2013a). For a more detailed introduction to protective measurement see Gao (2013b).

$[H_D, PA]$ is zero or its effect is negligible in the interval $[0, T]$ for simplicity. Thus in the limit of $N \to \infty$, we have

$$|t = T\rangle = |\psi\rangle\, e^{-\frac{i}{\hbar}\int_0^T g(t)\langle\psi|A|\psi\rangle P dt}\, |\phi(T)\rangle = |\psi\rangle\, e^{-\frac{i}{\hbar}<A>P}\, |\phi(T)\rangle\,, \tag{3}$$

where $< A >\equiv \langle\psi|A|\psi\rangle$ is the expectation value of the measured observable A in the measured state ψ. Since the total probability of other branches is proportional to T^2/N to the first order of N, the above state will be the state of the combined system after T when $N \to \infty$[2]. It can be seen that the exponential operator in Eq. (3) shifts the center of the pointer by an amount $\langle A\rangle$, namely that the state of the pointer after T is $\langle x + \langle A\rangle\, |\phi(T)\rangle$. This indicates that the result of the protective measurement is the expectation value of the measured observable in the measured state, and moreover, the measured state is not changed by the protective measurement.

It is worth stressing that under the above protection procedure the measurement of an observable is not necessarily weak (when compared with the system's Hamiltonian), and the measurement time T is not necessarily long enough so that the measurement interaction is adiabatic either (cf. Vaidman 2008)[3]. A stronger measurement with a shorter measurement time is better if only the projective measurements can be made frequently during the measurement (i.e. the condition $N \to \infty$ can be approximately valid). Certainly, the measurement must be weaker than the projective measurements which protect the measured state from being changed.

To sum up, we have demonstrated that for an arbitrary but known state of a quantum system at a given instant, we can protect the state from being changed via the quantum Zeno effect by frequent projective measurements, and an independent measurement of an observable A, which is made at the same time, yields the expectation value of the observable in the measured state[4].

By a conventional measurement on a single quantum system, one obtains one of the eigenvalues of the measured observable, and the expectation value of the observable can only be obtained as the statistical average of eigenvalues for an ensemble of identically prepared systems. Thus it seems surprising that a protective measurement can yield the expectation value of the measured observable only from a single quantum system. In fact, the appearance of expectation values as measurement results is quite natural when the measured state is not changed and the entanglement during the conventional measurement does not take place as for protective measurements[5]. In this case, the evolution of the

[2]Note that this result, like the quantum Zeno effect, does not depend on a particular formulation of quantum mechanics, and especially, it is independent of whether wavefunction collapse is real or apparent.

[3]According to Vaidman (2008), "Apart from protection, the procedure consists of a standard von Neumann measurement with weak coupling which is switched on and, after a long time, switched off, adiabatically."

[4]As we will argue later, in order to analyze the physical meaning of the wave function, we need not to measure the time evolution of the wave function, but only to measure the wave function at a given instant.

[5]The measured state being unchanged permits the state as well as the expectation values of observables in the state to be measurable. In this sense, protective measurement is not special; it is just the *very* way to measure the actual state of a quantum system at a given instant. By comparison, a non-protective measurement such as a conventional impulsive measurement will change the measured state, and the resulting measurement result (i.e. one of the eigenvalues of

combining state is

$$|\psi(0)\rangle |\phi(0)\rangle \rightarrow |\psi(t)\rangle |\phi(t)\rangle , t > 0 \tag{4}$$

where $|\psi(t)\rangle$ is the same as $|\psi(0)\rangle$ up to a phase factor during the measurement time interval $[0, T]$. Then by Ehrenfest's theorem we have

$$\frac{d}{dt}\langle\psi(t)\phi(t)|X |\psi(t)\phi(t)\rangle = -g(t)\langle\psi(0)|A |\psi(0)\rangle , \tag{5}$$

which further leads to

$$\langle\psi(T)\phi(T)|X |\psi(T)\phi(T)\rangle - \langle\psi(0)\phi(0)|X |\psi(0)\phi(0)\rangle = \langle\psi(0)|A |\psi(0)\rangle . \tag{6}$$

This means that the shift of the center of the pointer of the device gives the expectation value of the measured observable in the measured state.

Note that in some special cases the universal protection procedure via the quantum Zeno effect is not necessary, and the system's Hamiltonian can help protect its state from changing when the measurement interaction is weak and adiabatic. For example, for a quantum system in a discrete nondegenerate energy eigenstate, the system itself supplies the protection of the state due to energy conservation. By the adiabatic theorem, the adiabatic interaction during the measurement ensures that the measured system cannot make a transition from one discrete energy eigenstate to another. Moreover, according to the first order perturbation theory, for any given value of P, the energy of the measured energy eigenstate shifts by an infinitesimal amount: $\delta E = \langle H_I\rangle = g(t)P\langle A\rangle$, and the corresponding time evolution $e^{-iP\langle A\rangle/\hbar}$ then shifts the pointer by the expectation value $\langle A\rangle$. For degenerate energy eigenstates, we may not use the universal protection procedure either. The simplest way is to add a protective potential to change the energies of the other states and lift the degeneracy. Then the measured state remains unchanged, but is now protected by energy conservation like nondegenerate energy eigenstates.

Since the wave function can be reconstructed from the expectation values of a sufficient number of observables, the wave function of a quantum system can be measured by a series of protective measurements. Let the explicit form of the measured state at a given instant t be $\psi(x)$, and the measured observable A be (normalized) projection operators on small spatial regions V_n having volume v_n:

$$A = \begin{cases} \frac{1}{v_n}, & \text{if } x \in V_n, \\ 0, & \text{if } x \notin V_n. \end{cases} \tag{7}$$

A protective measurement of A then yields

$$\langle A\rangle = \frac{1}{v_n}\int_{V_n} |\psi(x)|^2 dv, \tag{8}$$

the measured observable) does not reflect the actual state of the measured system. Moreover, when a quantum system interacts with another quantum system under general non-protective conditions, its state also evolves in time, and thus the expectation values of observables do not manifest themselves explicitly in the interaction either. For example, the interaction between two charged quantum systems are not directly dependent on the expectation values of their charges, but described by the potential terms in the Schrödinger equation.

which is the average of the density $\rho(x) = |\psi(x)|^2$ over the small region V_n. Similarly, we can measure another observable $B = \frac{\hbar}{2mi}(A\nabla + \nabla A)$. The measurement yields

$$\langle B\rangle = \frac{1}{v_n}\int_{V_n}\frac{\hbar}{2mi}(\psi^*\nabla\psi - \psi\nabla\psi^*)dv = \frac{1}{v_n}\int_{V_n} j(x)dv. \tag{9}$$

This is the average value of the flux density $j(x)$ in the region V_n. Then when $v_n \to 0$ and after performing measurements in sufficiently many regions V_n we can measure $\rho(x)$ and $j(x)$ everywhere in space. Since the wave function $\psi(x,t)$ can be uniquely expressed by $\rho(x,t)$ and $j(x,t)$ (except for an overall phase factor), the whole wave function of the measured system at a given instant can be measured by protective measurements.

We have been discussing the protective measurement of a single quantum system. The scheme of protective measurement can also be extended to a many-particle system (Anandan 1993). If the system is in a product state, then one can easily measuring each state of the individual systems protectively. If the system is in an entangled state, one need to add an appropriate protection procedure to the whole system, and then the entangled state of the system can be protectively measured. However, it is worth noting that the realization of such protective measurements relies on the availability of corresponding protective interactions, which is limited by existing physical interactions[6].

Lastly, we stress that the validity of the scheme of protective measurement does not rely on the standard von Neumann formulation of measurement, in which it might be debatable to represent a macroscopic device with a single wave function. In the above formulation of protective measurement, the measuring system can be a microscopic system such as an electron, and the shift of the center of the wave packet of the measuring system is only determined by the Schrödinger equation. Since the state of the measured system is not changed during the protective measurement, a large number of identical measuring systems can be used to protectively measure the original measured system, and the centers of their wave packets have the same shift. Then the shift can be read out by conventional measurements of the ensemble of these identical measuring systems, for which the probability distribution of the results satisfies the Born rule. In a word, the scheme of protective measurement is only based on the Schrödinger equation (for microscopic systems) and the Born rule, and especially, it is independent of whether wavefunction collapse is real or not.

3 On the reality of the wave function

What are the physical implications of protective measurements? Several authors, including the inventors of protective measurements, have given some analyses of this question (Aharonov and Vaidman 1993; Aharonov, Anandan and Vaidman 1993; Anandan 1993; Dickson 1995). According to Aharonov and Vaidman (1993), protective measurement shows that the expectation values of observables are properties of a single quantum system. Moreover, it provides a strong argument for associating physical reality with the wave function of a

[6]For example, it seems that the entangled spatial wave function of a many-body system, which lives on configuration space, cannot be protectively measured. The reason is that a protective measurement cannot be performed in different positions in space at the same time.

single system. In particular, they thought that the wave function describes an extended object or a real physical wave. According to Anandan (1993), protective measurement refutes an argument of Einstein in favor of the ensemble interpretation of quantum mechanics. Dickson's (1995) analysis was more philosophical. He argued that protective measurement provides a reply to scientific empiricism about quantum mechanics, but it can neither refute that position nor confirm scientific realism, and the aim of his argument is to place realism and empiricism on an even score in regards to quantum mechanics. It seems that all these arguments rely on the presupposition that protective measurements are completely reliable (see, e.g. Vaidman 2009). This presupposition was objected notably by Dass and Qureshi (1999), as a realistic protective measurement can never be performed on a single quantum system with absolute certainty. For example, for a realistic protective measurement of an observable A in a non-degenerate energy eigenstate whose measuring interval T is finite, there is always a tiny probability proportional to $1/T^2$ to obtain a different result $\langle A\rangle_\perp$, where $\perp$ refers to a normalized state in the subspace normal to the measured state as picked out by the first order perturbation theory[7]. Our following arguments will avoid this major objection.

According to quantum mechanics, we can prepare a single measured system whose wave function is $\psi(t)$ at a given instant t. For example, the measured system is an electron being in the ground state of a Hydrogen atom. Now, by a protective measurement, we can obtain the expectation value of the measured observable in this state without disturbing the state (though with probability smaller than one in realistic situations)[8]. Moreover, by a series of protective measurements of certain observables, we can obtain the value of $\psi(t)$ *only* from this measured system. Thus we can reach the conclusion that the expectation values of observables are the physical properties of a single quantum system, and the wave function of the system represents the physical property of the system[9]. In particular, $\psi(x,t)$, the spatial wave function of the system in position x at instant t, represents the physical property of the system in position x at instant t. This also means that for a quantum system, there is a physical entity spreading out over a region of space where the spatial wave function of the system is not zero.

Here we assume a realist view on the theory-reality relation, which means that the theoretical terms expressed in the language of mathematics connect to

[7]Moreover, after obtaining the result $\langle A\rangle_\perp$ the measured state also collapses to the state $\perp$ according to the standard formulation of quantum mechanics.

[8]When the measurement obtains the expectation value of the measured observable in the measured state, the measured state is not changed. Moreover, the probability of obtaining a different result and collapsing the measured state can be made arbitrarily small in principle. By comparison, the eigenvalues values of the measured observable being measurement results are only consequences of non-protective, strong measurements, which disturb the measured state strongly and are arguably not good, qualified measurements.

[9]There might also exist other components of the underlying physical state, which are not measureable by protective measurements and not described by the wave function, e.g. the positions of the Bohmian particles in the de Broglie-Bohm theory (de Broglie 1928; Bohm 1952). In this case, however, the wave function is still uniquely determined by the underlying physical state, though it is not a complete representation of the physical state. As a result, the epistemic interpretation of the wave function will be ruled out (cf. Lewis et al 2012). Certainly, the wave function also plays an epistemic role by giving the probability distribution of the results of conventional impulsive measurements according to the Born rule. However, this role is secondary and determined by the complete quantum dynamics that describes the measuring process, e.g. the collapse dynamics in dynamical collapse theories.

the entities existing in the physical world. On this view, the wave function in quantum mechanics describes either the state of an ensemble of identical systems or the state of a single system. Since we can measure the wave function only from a single system by protective measurements, the wave function must represent the property of a single system[10]. Note that this conclusion is independent of whether the wave function of the measured system is known beforehand for protective measurements. Even though we know the wave function, which is an abstract mathematical object, we still don't know its physical meaning, while protective measurements can help answer this fundamental question of quantum mechanics[11].

4 Meaning of the wave function

In this section, we will further analyze the physical state described by the wave function and measurable by protective measurements. As we will see, the analysis may provide an important clue to the meaning of the wave function[12].

4.1 The existence of effective charge distribution

As we have shown in Section 2, the result of a protective measurement is not directly the wave function, but the density or flux density. What density? When the interaction used to measure the observable A defined in Eq.(7) is physically realized by electromagnetic interaction between the measured system and the measuring system, the density measured by the protective measurement will be the effective charge density of the system in the measured position[13]. During

[10]We can also give a PBR-like argument for ψ-ontology in terms of protective measurements (cf. Pusey, Barrett and Rudolph 2012). For two (known) nonorthogonal states of a quantum system, the results of the protective measurements of them may be different with probability that can be arbitrarily close to one. If there exists a finite probability that these two nonorthogonal states correspond to the same physical state λ, then when assuming λ determines the probability of measurement results as the PBR theorem assumes, the results for the two nonorthogonal states will be the same with the finite probability. This leads to a contradiction. This argument only considers one quantum system, and avoids the independence assumption used by the PBR theorem.

[11]In addition, as pointed out by Aharonov, Anandan and Vaidman (1996), the wave function of the measured system may be unknown beforehand when splitting the procedure of a protective measurement into two stages. The first is a protection, made by one experimenter or even just by nature, and the second is performed by another experimenter who does not know the measured state. What this experimenter needs to know is that the state is protected and what is the degree of protection, and he does obtain new information by protective measurement.

[12]Quantum mechanics is a physical theory about the wave function and its evolution. The most fundamental interpretative problem of the theory is the physical meaning of the wave function. Unfortunately, it has been treated as a marginal problem, especially compared with the measurement problem (however, see Ney and Albert 2013 for a recent exception). There are already several alternatives to quantum mechanics which give respective solutions to the measurement problem, such as the de Broglie-Bohm theory and the many-worlds interpretation (de Broglie 1928; Bohm 1952; Everett 1957; De Witt and Graham 1973). However, these theories at their present stages are unsatisfactory at least in one aspect; they have not succeeded in making sense of the wave function. In the following, we will argue that this fundamental interpretative problem may be solved independently of how to solve the measurement problem.

[13]Similarly, protective measurements can also measure the effective electric flux density of the system. This important point was also admitted by Aharonov and Vaidman (1993).

the measurement, the wave function of the measuring system, $\phi(x,t)$, will obey the following Schrödinger equation:

$$i\hbar\frac{\partial\phi(x,t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\phi(x,t)-k\frac{e\cdot|\psi(x_s,t)|^2dx_sQ}{|x-x_s|}\phi(x,t), \tag{10}$$

where m is the mass of the measuring system, k is the Coulomb constant, x_s is the measured position, Q is the charge of the measured system, and $\psi(x,t)$ is the wave function of the system. From this equation, it can be seen that the measured quantity or the property of the measured system in the measured position x_s that has efficiency to influence the measuring system is $|\psi(x_s,t)|^2dx_sQ$, the effective charge there, and when divided by the volume element dx_s it is the effective charge density $|\psi(x_s,t)|^2Q$[14]. This means that the charge of a quantum system is distributed throughout space in efficiency, and the effective charge density in each position is proportional to the modulus squared of the wave function of the system there[15].

4.2 The origin of effective charge distribution

What entity or process, then, generates the physical efficiency of the quantity $|\psi(x,t)|^2dxQ$ or the effective charge distribution in space?[16] It can be expected that the answer will help understand the meaning of $|\psi(x,t)|^2$ and the wave function $\psi(x,t)$ itself. There are two possibilities: the effective charge distribution can be generated by either (1) a continuous charge distribution with density $|\psi(x,t)|^2Q$ or (2) the motion of a discrete point charge Q with spending time $|\psi(x,t)|^2dxdt$ in each infinitesimal spatial volume $[x,x+dx]$ in the infinitesimal time interval $[t,t+dt]$. For the first possibility, the charge distribution exists

[14]For a two-body system, the effective charge density in each position x is $\rho_Q(x,t)=Q_1\int_{-\infty}^{+\infty}|\psi(x,x_2,t)|^2dx_2+Q_2\int_{-\infty}^{+\infty}|\psi(x_1,x,t)|^2dx_1$, where Q_1 and Q_2 are the charges of the two subsystems, respectively, and $\psi(x_1,x_2,t)$ is the wave function of the whole system. The existence of the effective charge distribution may also be seen from the potential terms in the Schrödinger equation more directly. For instance, the electrostatic interaction term $Q\varphi\psi(x,t)$ in the Schrödinger equation for a charged quantum system indicates that the electrostatic interaction exists in all regions where $\psi(x,t)$ is nonzero, where Q is the charge of the system, $\psi(x,t)$ is the wave function of the system, and φ is an external electric scalar potential. Thus the charge of the system should also distribute throughout these regions. If the charge did not distribute and have efficiency in regions where the wave function is nonzero, then there would not exist electrostatic interaction there. Note that this argument assumes that the wave function represents the physical state of a single quantum system. Historically, the charge density interpretation for electrons was originally suggested by Schrödinger in his fourth paper on wave mechanics (Schrödinger 1926). Schrödinger clearly realized that the charge density cannot be classical because his equation does not include the usual classical interaction between the densities. Presumably since people thought that the charge density could not be measured and also lacked a consistent physical picture, this interpretation was soon rejected and replaced by Born's probability interpretation. Now protective measurement re-endows the effective charge distribution of an electron with reality by a more convincing argument. The question is then how to find a consistent physical explanation for it. Our following analysis may be regarded as a further development of Schrödinger's original idea to some extent. For more discussions of Schrödinger's charge density interpretation see Bacciagaluppi and Valentini (2009).

[15]Similarly, the mass of a quantum system is also distributed throughout space, and the mass density in each position is proportional to the modulus squared of the wave function of the system there.

[16]Note again that the efficiency is not manifested under special conditions; rather, it is just the manifestation of the wave function itself, which is protected to be not changed during the interaction.

throughout space at the same time, while for the second possibility, at every instant there is only a localized, point-like particle with the total charge of the system, and its motion during an infinitesimal time interval forms the effective charge distribution. Concretely speaking, at a particular instant the charge density of the particle in each position is either zero (if the particle is not there) or singular (if the particle is there), while the time average of the density during an infinitesimal time interval gives the effective charge density. Moreover, the motion of the particle is ergodic in the sense that the integral of the formed charge density in any region is required to be equal to the expectation value of the total charge in the region. We will give two arguments supporting the second possibility below.

First of all, it can be argued that the effective charge distribution cannot be generated by a continuous charge distribution for many-body systems. For a one-body system, a continuous charge distribution, which is described by charge density and electric flux density, may represent the physical state of the system described by its wave function, as there exists a one-to-one corresponding relation between density, flux density and the wave function (except for an overall phase factor). However, for a N-body system, a continuous charge distribution cannot represent the physical state of the system described by its wave function, as the density and flux density of N sub-systems defined in 3-dimensional space are not enough to consititute the wave function defined in 3N-dimensional configuration space[17]. In other words, the continuous charge distribution does not contain the information about the entanglement between the sub-systems of the many-body system[18]. By comparison, as we will see later, the motion of particles may contain the entanglement information, and the state of motion of N particles in 3-dimensional space may be described by the wave function in 3N-dimensional space.

Next, it can be argued that the existence of a continuous charge distribution for a one-body system may also lead to inconsistency. If the charge distribution is continuous and exists throughout space at the same time, then any two parts of the distribution, like two electrons, will arguably have electrostatic interaction described by the interaction potential term in the Schrödinger equation. However, the existence of such electrostatic self-interaction for a quantum system contradicts the superposition principle of quantum mechanics (at least for microscopic systems such as electrons). Moreover, the existence of the electrostatic self-interaction for the effective charge distribution of an electron is incompati-

[17]It is obvious that a continuous charge distribution in a 3N-dimensional configuration space cannot represent the physical state of a N-body system either. For example, two charges Q1 and Q2 being in two positions in 3-dimensional space cannot be represented by a charge Q in one position in a 6-dimensional configuration space.

[18]Note that the 3N coordinates of a point in a configuration space are N groups of three position coordinates in our three-dimensional space. Here is a simple argument. Suppose the wave function of a two-body system is localized in a point $(x_1, y_1, z_1, x_2, y_2, z_2)$ in the configuration space of the system. We make a position measurement on the system. Quantum mechanics predicts that only in positions (x_1, y_1, z_1) and (x_2, y_2, z_2) in our three-dimensional space can we obtain a measurement result. This means that if the wave function represents the property or state of a physical entity, then this entity exists in positions (x_1, y_1, z_1) and (x_2, y_2, z_2) in the three-dimensional space. This explains the meaning of the coordinates of a point in the configuration space where the wave function lives. Note that this argument is independent of whether our three-dimensional space is fundamental or emergent from the 3N-dimensional configuration space. For more recent analyses of configuration space and wavefunction realism see Monton (2002) and Lewis (2004).

ble with experimental observations either. For example, for the electron in the 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 would be remarkably different from those predicted by quantum mechanics and confirmed by experiments if there existed such electrostatic self-interaction. By contrast, if there is only a localized particle at every instant, it is understandable that there exists no electrostatic self-interaction of the effective charge distribution formed by the motion of the particle. This is consistent with the superposition principle of quantum mechanics and experimental observations[19].

To sum up, we have argued that the effective charge distribution of a quantum system such as an electron originates from the ergodic motion of a discrete point charge. At every instant there is only a localized particle with the total mass and charge of the system, while during an infinitesimal time interval the ergodic motion of the particle forms the effective mass and charge distribution. The spending time of the particle in each position is proportional to the modulus squared of the wave function of the system there, and the formed mass and charge density in each position is equal to the modulus squared of the wave function of the system there multiplied by the total mass and charge of the system, respectively.

4.3 Ergodic motion of particles

Which sort of ergodic motion? This is a further question that needs to be answered. If the ergodic motion of a particle is continuous, then it can only form

[19]Here is a further clarification of this argument. It can be seen that the non-existence of self-interaction of the charge distribution poses a puzzle. According to quantum mechanics, two charge distributions such as two electrons, which exist in space at the same time, have electrostatic interaction described by the interaction potential term in the Schrödinger equation, but for the effective charge distribution of an electron, any two parts of the distribution have no such electrostatic interaction. Facing this puzzle one may have two choices. The first one is simply admitting that the non-existence of self-interaction of the effective charge distribution is a distinct feature of the laws of quantum mechanics, but insisting that the laws are what they are and no further explanation is needed. However, this choice seems to beg the question and is unsatisfactory in the final analysis. A more reasonable choice is to try to explain this puzzling feature, e.g. by analyzing its relationship with the existent form of the effective charge distribution. The effective charge distribution has two possible origins or forms after all. On the one hand, the non-existence of self-interaction of the distribution may help determine which possible form is the actual one. For example, one possible form is inconsistent with this distinct feature, while the other possible form is consistent with it. On the other hand, the actual existent form of the effective charge distribution may also help explain the non-existence of self-interaction of the distribution. This is just what the above argument has done. The analysis establishes a connection between the non-existence of self-interaction of the effective charge distribution and the actual existent form of the distribution. The reason why two wavepackets of an electron, each of which has part of the electron's charge in efficiency, have no electrostatic interaction is that these two wavepackets do not exist at the same time, and their effective charges are formed by the motion of a localized particle with the total charge of the electron. Since there is only a localized particle at every instant, it is understandable that there exists no electrostatic self-interaction of the effective charge distribution formed by the motion of the particle. By contrast, if the two wavepackets with charges, like two electrons, existed at the same time, then they would also have the same form of electrostatic interaction as that between two electrons. Note that this argument does not assume that real charges which exist at the same time are classical charges and they have classical interaction. By contrast, the Schrödinger-Newton equation, which was proposed by Diósi (1984) and Penrose (1998), treats the mass distribution of a quantum system as classical.

the effective mass and charge distributions during a finite time interval[20]. But according to quantum mechanics, the effective mass and charge distributions at a given instant are required to be formed by the ergodic motion of the particle during an infinitesimal time interval near the instant. Thus it seems that the ergodic motion of the particle cannot be continuous but must be discontinuous. This is at least what the existing theory says. This conclusion can also be reached by analyzing a specific example. Consider an electron in a superposition of two energy eigenstates in two separate boxes. In this example, even if one assumes that the electron can move with infinite velocity, it cannot *continuously* move from one box to another due to the restriction of box walls. Therefore, any sort of continuous motion cannot generate the effective charge distribution that exists in both boxes[21].

On the other hand, in order that the ergodic motion of a particle forms the right mass and charge distributions, for which the mass and charge density in each position is proportional to the modulus squared of its wave function there, the (objective) probability density for the particle to appear in each position must be proportional to the modulus squared of its wave function there too (and for normalized wave functions they should be equal)[22]. This is understandable, since that the mass and charge density is large in a position requires that the spending time of the particle is long there or the frequency of the particle appearing there is high. Moreover, from a logical point of view, the particle must also have an instantaneous property (as a probabilistic instantaneous condition) which determines the probability density for it to appear in every position in space; otherwise the particle would not "know" how frequently it should appear in each position in space. This property is usually called indeterministic disposition or propensity in the literature[23].

In conclusion, we have argued that the consistency of the formed mass and charge distribution with that predicted by quantum mechanics requires that the ergodic motion of a particle is discontinuous, and the probability density for the particle to appear in every position is equal to the modulus squared of its wave function there. In other words, the ergodic motion of the particle is random and discontinuous.

[20]For other objections to classical ergodic models see Aharonov and Vaidman (1993) and Aharonov, Anandan and Vaidman (1993).

[21]One may object that this is merely an artifact of the idealization of infinite potential. However, even in this ideal situation, the ergodic model should also be able to generate the effective charge distribution by means of some sort of ergodic motion of the electron; otherwise it will be inconsistent with quantum mechanics.

[22]Besides, for normalized wave functions, the (objective) probability current density must also equal to the formed mass or charge flux density divided by the mass or charge of the particle.

[23]Note that the propensity here denotes single case propensity. Long run propensity theories fail to explain objective single-case probabilities. According to these theories, it makes no sense to speak of the propensity of a single isolated event in the absence of a sequence that contains it. In addition, it is worth stressing that the propensities possessed by particles relate to their objective motion, not to the measurements on them. By contrast, according to the existing propensity interpretations of quantum mechanics, the propensities a quantum system has relate only to measurements; a quantum system possesses the propensity to exhibit a particular value of an observable if the observable is measured on the system. These interpretations cannot be wholly satisfactory because of resorting to the vague concept of measurement.

4.4 Interpreting the wave function

According to the above analysis, microscopic particles such as electrons, which are described by quantum mechanics, are indeed particles. Here the concept of particle is used in its usual sense. A particle is a small localized object with mass and charge, and it is only in one position in space at an instant. Moreover, the motion of these particles is not continuous but discontinuous and random in nature. We may say that an electron is a quantum particle in the sense that its motion is not continuous motion described by classical mechanics, but random discontinuous motion described by quantum mechanics.

Unlike the deterministic continuous motion, the trajectory function $x(t)$ can no longer provide a useful description for random discontinuous motion. It has been shown that the strict description of random discontinuous motion of a particle can be given based on the measure theory (Gao 2013b). Loosely speaking, the random discontinuous motion of the particle forms a particle "cloud" extending throughout space (during an infinitesimal time interval), and the state of motion of the particle is represented by the density and flux density of the cloud, denoted by $\rho(x,t)$ and $j(x,t)$, respectively, which satisfy the continuity equation $\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial j(x,t)}{\partial x} = 0$. The density of the cloud, $\rho(x,t)$, represents the probability density for the particle to appear in position x at instant t, and it satisfies the normalization condition $\int_{-\infty}^{+\infty} \rho(x,t)dx = 1$.

As we have argued above, for a charged particle such as an electron, the cloud will be an electric cloud, and $\rho(x,t)$ and $j(x,t)$, when multiplied by the total charge of the particle, will be the (effective) charge density and electric flux density measurable by protective measurements, respectively. Thus we have the following relations:

$$\rho(x,t) = |\psi(x,t)|^2, \tag{11}$$

$$j(x,t) = \frac{\hbar}{2mi}[\psi^*(x,t)\frac{\partial \psi(x,t)}{\partial x} - \psi(x,t)\frac{\partial \psi^*(x,t)}{\partial x}]. \tag{12}$$

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

$$\psi(x,t) = \sqrt{\rho(x,t)}e^{im\int_{-\infty}^{x} \frac{j(x',t)}{\rho(x',t)}dx'/\hbar}. \tag{13}$$

This means that the wave function $\psi(x,t)$ also provides a description of the state of random discontinuous motion of a particle.

This picture of motion of a single particle can be extended to the motion of many particles. The extension may explain the multi-dimensionality of the wave function (cf. Monton 2002; Lewis 2004). At a given instant, a quantum system of N particles can be represented by a point in a $3N$-dimensional configuration space. During an infinitesimal time interval, these particles perform random discontinuous motion in the real space, and correspondingly, this point performs random discontinuous motion in the configuration space and forms a cloud there. Then, similar to the single particle case, the state of the system is represented by the density and flux density of the cloud in the configuration space, $\rho(x_1, x_2, ...x_N, t)$ and $j(x_1, x_2, ...x_N, t)$, where the density $\rho(x_1, x_2, ...x_N, t)$ represents the probability density that particle 1 appears in position x_1 and particle 2 appears in position x_2, ..., and particle N appears in

position x_N[24]. Since these two quantities are defined in the 3N-dimensional configuration space, the many-particle wave function, which is composed of these two quantities, is also defined in the 3N-dimensional configuration space.

One important point needs to be stressed here. Since the wave function in quantum mechanics is defined at a given instant, not during an infinitesimal time interval, it should be regarded not simply as a description of the state of motion of particles, but more suitably as a description of the dispositional property of the particles that determines their random discontinuous motion at a deeper level[25]. In particular, the modulus squared of the wave function determines the probability density that the particles appear in certain positions in space. By contrast, the density and flux density of the particle cloud in configuration space, 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.

4.5 On momentum, energy and spin

We have been discussing random discontinuous motion of particles in real space. Does the picture of random discontinuous motion exist for other dynamical variables such as momentum and energy? Since there are also wave functions of these variables in quantum mechanics, it seems tempting to assume that the above interpretation of the wave function in position space also applies to the wave functions in momentum space etc[26]. This means that when a particle is in a superposition of the eigenstates of a variable, it also undergoes random discontinuous motion among the corresponding eigenvalues of this variable. For example, a particle in a superposition of energy eigenstates also undergoes random discontinuous motion among all energy eigenvalues. At each instant the energy of the particle is definite, randomly assuming one of the energy eigenvalues with probability given by the modulus squared of the wave function at this energy eigenvalue, and during an infinitesimal time interval the energy of the particle spreads throughout all energy eigenvalues. Since the values of two non-commutative variables (e.g. position and momentum) at every instant may be mutually independent, the objective value distribution of every variable can be equal to the modulus squared of its wave function and consistent with quantum mechanics [27].

[24]When these N particles are independent, the density $\rho(x_1, x_2, ...x_N, t)$ can be reduced to the direct product of the density for each particle, namely $\rho(x_1, x_2, ...x_N, t) = \prod_{i=1}^{N} \rho(x_i, t)$.

[25]For a many-particle system in an entangled state, this dispositional property is possessed by the whole system.

[26]Under this assumption, the ontology of the theory will not only include the wavefunction and the particle position, but also include momentum and energy.

[27]Note that for random discontinuous motion a property (e.g. position) of a quantum system in a superposed state of the property is indeterminate in the sense of usual hidden variables, though it does have a definite value at each instant. For this reason, the particle position should not be called hidden variable for random discontinuous motion of particles, and the resulting theory is not a hidden variable theory either. This makes the theorems that restrict hidden variables such as the Kochen-Specker theorem irrelevant. Another way to see this is to realize that random discontinuous motion of particles alone does not provide a way to solve the measurement problem, and wavefunction collapse may also be needed. For details see Gao (2013b).

However, there is also another possibility, namely that the picture of random discontinuous motion exists only for position, while momentum, energy etc do not undergo random discontinuous change among their eigenvalues. This is a minimum formulation in the sense that the ontology of the theory only includes the wave function and the particle position. On this view, the position of a particle is an instantaneous property of the particle defined at instants, while momentum and energy are properties relating only to its state of motion (e.g. momentum and energy eigenstates), which is formed by the motion of the particle during an infinitesimal time interval[28]. This may avoid the problem of defining the momentum and energy of a particle at instants. Certainly, we can still talk about momentum and energy on this view. For example, when a particle is in an eigenstate of the momentum or energy operator, we can say that the particle has definite momentum or energy, whose value is the corresponding eigenvalue. Moreover, when a particle is in a momentum or energy superposition state and the momentum or energy branches are well separated in space, we can still say that the particle has definite momentum or energy in certain local regions.

Lastly, we note that spin is a more distinct property. Since the spin of a free particle is always definite along one direction, the spin of the particle does not undergo random discontinuous motion, though a spin eigenstate along one direction can always be decomposed into two different spin eigenstates along another direction. But if the spin state of a particle is entangled with its spatial state due to interaction and the branches of the entangled state are well separated in space, the particle in different branches will have different spin, and it will also undergo random discontinuous motion between these different spin states. This is the situation that usually happens during a spin measurement.

5 Conclusions

Protective measurement is a measuring method, by which one can measure the expectation value of an observable on a single quantum system, even if the system is initially not in an eigenstate of the measured observable. This remarkable feature makes protective measurements quite distinct from conventional impulsive measurements, and as we have argued above, it may have important implications for our understandings of quantum mechanics.

According to the standard view, the expectation values of observables are not the physical properties of a single system, but the statistical properties of an ensemble of identical systems. This seems reasonable if there exist only conventional impulsive measurements. An impulsive measurement can only obtain one of the eigenvalues of the measured observable, and thus the expectation value can only be defined as a statistical average of the eigenvalues for an ensemble of identical systems. However, there exist other kinds of quantum measurements, and in particular, protective measurements can measure the expectation values of observables for a single system. Therefore, the expectation values of observables should be taken as the physical properties of a single quantum system. This is the first conceptual shift brought by protective measurement.

[28]It is worth stressing that the particle position here is different from the position property described by the position operator in quantum mechanics, and the latter is also a property relating only to the state of motion of the particle such as position eigenstates.

Since the wave function can be reconstructed from the expectation values of a sufficient number of observables, the above result will immediately lead to the second implication, namely that the wave function of a quantum system is a representation of the physical state (or ontic state) of the system. This result is more definite than that obtained by Pusey, Barrett and Rudolph (2012), which was based on an analysis of results of impulsive measurements and the independence assumption.

The more important virtue of protective measurement is that it can further help reveal the physical state represented by the wave function. Indeed, it provides for the first time a method to measure the actual physical state of a single quantum system. When a protective measurement is realized by electromagnetic or gravitational interaction between the measured system and the measuring device, it can measure the effective charge or mass distribution of the system. The results as predicted by quantum mechanics show that the mass and charge of a quantum system such as an electron is distributed throughout space in efficiency, and the effective mass and charge density in each position is proportional to the modulus squared of its wave function there.

The effective mass and charge distribution of a quantum system has two possible existent forms, and a further analysis is needed to find which one is the actual form. It can be argued that the superposition principle of quantum mechanics and the existence of quantum entanglement require that the effective mass and charge distribution is formed by the ergodic motion of a localized particle with the total mass and charge of the system. Moreover, the consistency of the formed distribution with that predicted by quantum mechanics requires that the ergodic motion of the particle is discontinuous, and the probability density that the particle appears in every position is equal to the modulus squared of its wave function there.

Therefore, according to the above analysis, it seems that quantum mechanics, like Newtonian mechanics, also deals with the motion of particles in space and time. Microscopic particles such as electrons are still particles, but they move in a discontinuous and random way. Moreover, the wave function describes the state of random discontinuous motion of particles, and at a deeper level, it represents the dispositional property of the particles that determines their random discontinuous motion. Quantum mechanics, in this way, is essentially a physical theory about the laws of random discontinuous motion of particles. It is a further and also harder question what the precise laws are, e.g. whether the wave function undergoes a stochastic and nonlinear collapse evolution [29].

Acknowledgments

I am very grateful to Dean Rickles, Huw Price, Guido Bacciagaluppi, David Miller, Vincent Lam, Antony Valentini, Hans Westman, and Lev Vaidman for helpful comments and suggestions. I am also grateful for stimulating discussions with participants in the PIAF (Perimeter Institute - Australia Foundations)

[29] It has been argued that protective measurement and the picture of random discontinuous motion of particles seem to suggest that wavefunction collapse is a real physical process, and dynamical collapse theories are in the right direction in solving the measurement problem (Gao 2013b). Note that the existence of very slow collapse of the wave function for microscopic systems does not influence the principle of protective measurement and its implications.

Workshop in Foundations, held at the Griffith University, Brisbane, December 1-3, 2010.

References

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

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

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

[4] Anandan, J. (1993). Protective measurement and quantum reality. Found. Phys. Lett., 6, 503-532.

[5] Bacciagaluppi, G. and Valentini, A. (2009). Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference. Cambridge: Cambridge University Press.

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

[7] Dass, N. D. H. and Qureshi, T. (1999). Critique of protective measurements. Phys. Rev. A 59, 2590.

[8] 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. See Bacciagaluppi and Valentini (2009) for an English translation.

[9] DeWitt, B. S. and Graham, N. (eds.) (1973). The Many-Worlds Interpretation of Quantum Mechanics, Princeton: Princeton University Press.

[10] Dickson, M. (1995). An empirical reply to empiricism: protective measurement opens the door for quantum realism. Philosophy of Science 62, 122.

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

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

[13] Gao, S. (2013a). On Uffink's criticism of protective measurements, Studies in History and Philosophy of Modern Physics (In Press). Available online 7 August 2013, http://dx.doi.org/10.1016/j.shpsb.2013.07.001.

[14] Gao, S. (2013b). Interpreting Quantum Mechanics in Terms of Random Discontinuous Motion of Particles. http://philsci-archive.pitt.edu/9589/.

[15] Ghose, P. and Home, D. (1995). An analysis of Aharonov-Anandan-Vaidman model, Found. Phys. 25, 1105.

[16] Lewis, P. (2004). Life in configuration space, British Journal for the Philosophy of Science 55, 713-729.

[17] Lewis, P. G., Jennings, D., Barrett, J., and Rudolph, T. (2012). Distinct quantum states can be compatible with a single state of reality. Phys. Rev. Lett. 109, 150404.

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

[19] Ney, A. and Albert, D. Z. (eds.) (2013). The Wave Function: Essays on the Metaphysics of Quantum Mechanics. Oxford: Oxford University Press.

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

[21] Pusey, M., Barrett, J. and Rudolph, T. (2012). On the reality of the quantum state. Nature Phys. 8, 475-478.

[22] Rovelli, C. (1994). Comment on "Meaning of the wave function", Phys. Rev. A 50, 2788.

[23] Schrödinger, E. (1926). Quantizierung als Eigenwertproblem (Vierte Mitteilung). Ann. d. Phys. (4) 81, 109-139. English translation: Quantisation as a Problem of Proper Values. Part IV, Reprint in Schrödinger, E. (1982). Collected Papers on Wave Mechanics. New York: Chelsea Publishing Company, pp. 102-123.

[24] Uffink, J. (1999). How to protect the interpretation of the wave function against protective measurements. Phys. Rev. A 60, 3474-3481.

[25] Unruh, W. G. (1994). Reality and measurement of the wave function Phys. Rev. A 50, 882.

[26] 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.