Protective Measurement: A Paradigm Shift in Understanding Quantum Mechanics

Shan Gao*

March 5, 2013

Abstract

This article introduces the method of protective measurement and discusses its deep implications for the foundations of quantum mechanics.

1 Introduction

Protective measurement, in the language of standard quantum mechanics, is a method to measure the expectation value of an arbitrary 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 impulsive measurement, the coupling interaction between the measured system and the measuring device is of short duration and strong. By contrast, protective measurement uses a weak and long duration coupling interaction and an appropriate procedure to protect the measured system from being disturbed. A general scheme is to let the measured system be in a nondegenerate eigenstate of the whole Hamiltonian using a suitable protective interaction (in some situations the protection is provided by the measured system itself), and then make the measurement adiabatically so that the state of the system neither changes nor becomes entangled with the measuring device appreciably. In this way, such protective measurements can measure the expectation values of observables on a single quantum system, and in particular, the physical state of the system, which is described by its wave function, can also be measured as expectation values of certain observables.

2 Mathematical formulation of protective measurement

As a typical example, we consider a quantum system in a discrete nondegenerate energy eigenstate $|E_n\rangle$. In this case, the system itself supplies the protection of the state due to energy conservation and no artificial protection is needed[1].

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

[1]As will be shown below, before the protective measurement we only need to know the measured state is a discrete nondegenerate energy eigenstate of the Hamiltonian of the system,

According to the standard von Neumann procedure, measuring an observable A in this state involves an interaction Hamiltonian

$$H_I = g(t)PA \tag{1}$$

coupling the measured system to an appropriate measuring device, where P is the momentum conjugate to the pointer variable X of an appropriate measuring device. The time-dependent coupling strength $g(t)$ is a smooth function normalized to $\int dtg(t) = 1$ during the interaction interval T, and $g(0) = g(T) = 0$. The initial state of the pointer at $t = 0$ is supposed to be $|\phi(x_0)\rangle$, which is a Gaussian wave packet of eigenstates of X with width w_0, centered around the eigenvalue x_0.

For a conventional impulsive measurement, the interaction H_I is of very short duration and so strong that it dominates the rest of the Hamiltonian (i.e. the effect of the free Hamiltonians of the measuring device and the measured system can be neglected). Then the state of the combined system at the end of the interaction can be written as

$$|t = T\rangle = e^{-\frac{i}{\hbar}PA} |E_n\rangle |\phi(x_0)\rangle . \tag{2}$$

By expanding $|E_n\rangle$ in the eigenstates of A, $|a_i\rangle$, we obtain

$$|t = T\rangle = \sum_i e^{-\frac{i}{\hbar}Pa_i} c_i |a_i\rangle |\phi(x_0)\rangle , \tag{3}$$

where c_i are the expansion coefficients. The exponential term shifts the center of the pointer by a_i:

$$|t = T\rangle = \sum_i c_i |a_i\rangle |\phi(x_0 + a_i)\rangle . \tag{4}$$

This is an entangled state, where the eigenstates of A with eigenvalues a_i get correlated to measuring device states in which the pointer is shifted by these values a_i. Then by the collapse postulate of standard quantum mechanics, the state will instantaneously and randomly collapse into one of its branches $|a_i\rangle |\phi(x_0 + a_i)\rangle$ with probability $|c_i|^2$. This means that the measurement result can only be one of the eigenvalues of measured observable A, say a_i, with a certain probability, say $|c_i|^2$. The expectation value of A is then obtained as the statistical average of eigenvalues for an ensemble of identically prepared systems, namely $\langle A\rangle = \sum_i |c_i|^2 a_i$.

Different from the conventional impulsive measurements, for which the interaction is very strong and almost instantaneous, protective measurements make use of the opposite limit where the interaction of the measuring device with the system is weak and adiabatic, and thus the free Hamiltonians cannot be neglected[2]. Let the Hamiltonian of the combined system be

$$H(t) = H_S + H_D + g(t)PA, \tag{5}$$

and we need not to know the measured state or the Hamiltonian of the system or the measured state is one of a known collection of energy eigenstates. In this case, by a conventional impulsive measurement we can only measure the energy of the system, and we cannot measure the expectation value of any other observable of the system (as well as the wave function of the system).

[2]For a more detailed derivation of protective measurement see Dass and Qureshi (1999).

where H_S and H_D are the free Hamiltonians of the measured system and the measuring device, respectively. The interaction lasts for a long time T, and $g(t)$ is very small and constant for the most part, and it goes to zero gradually before and after the interaction.

The state of the combined system after T is given by

$$|t = T\rangle = e^{-\frac{i}{\hbar}\int_0^T H(t)dt} |E_n\rangle |\phi(x_0)\rangle . \tag{6}$$

By ignoring the switching on and switching off processes[3], the full Hamiltonian (with $g(t) = 1/T$) is time-independent and no time-ordering is needed. Then we obtain

$$|t = T\rangle = e^{-\frac{i}{\hbar}HT} |E_n\rangle |\phi(x_0)\rangle , \tag{7}$$

where $H = H_S + H_D + \frac{PA}{T}$. We further expand $|\phi(x_0)\rangle$ in the eigenstate of H_D, $\left|E_j^d\right\rangle$, and write

$$|t = T\rangle = e^{-\frac{i}{\hbar}HT} \sum_j c_j |E_n\rangle \left|E_j^d\right\rangle , \tag{8}$$

Let the exact eigenstates of H be $|\Psi_{k,m}\rangle$ and the corresponding eigenvalues be $E(k,m)$, we have

$$|t = T\rangle = \sum_j c_j \sum_{k,m} e^{-\frac{i}{\hbar}E(k,m)T} \langle \Psi_{k,m}|E_n, E_j^d\rangle |\Psi_{k,m}\rangle. \tag{9}$$

Since the interaction is very weak, the Hamiltonian H of Eq.(5) can be regarded as $H_0 = H_S + H_D$ perturbed by $\frac{PA}{T}$. Using the fact that $\frac{PA}{T}$ is a small perturbation and that the eigenstates of H_0 are of the form $|E_k\rangle \left|E_m^d\right\rangle$, the perturbation theory gives

$$\begin{aligned} |\Psi_{k,m}\rangle &= |E_k\rangle \left|E_m^d\right\rangle + O(1/T), \\ E(k,m) &= E_k + E_m^d + \frac{1}{T}\langle A\rangle_k \langle P\rangle_m + O(1/T^2). \end{aligned} \tag{10}$$

Substituting Eq.(10) in Eq.(9) and taking the limit $T \to \infty$ yields

$$|t = T\rangle_{T\to\infty} = \sum_j e^{-\frac{i}{\hbar}(E_n T + E_j^d T + \langle A\rangle_n \langle P\rangle_j)} c_j |E_n\rangle \left|E_j^d\right\rangle . \tag{11}$$

For the case where P commutes with the free Hamiltonian of the device[4], i.e., $[P, H_D] = 0$, the eigenstates $\left|E_j^d\right\rangle$ of H_D are also the eigenstates of P, and thus the above equation can be rewritten as

$$|t = T\rangle_{T\to\infty} = e^{-\frac{i}{\hbar}E_n T - \frac{i}{\hbar}H_D T - \frac{i}{\hbar}\langle A\rangle_n P} |E_n\rangle |\phi(x_0)\rangle . \tag{12}$$

It can be seen that the third term in the exponent will shift the center of the pointer $|\phi(x_0)\rangle$ by an amount $\langle A\rangle_n$:

$$|t = T\rangle_{T\to\infty} = e^{-\frac{i}{\hbar}E_n T - \frac{i}{\hbar}H_D T} |E_n\rangle |\phi(x_0 + \langle A\rangle_n)\rangle. \tag{13}$$

[3]The change in the total Hamiltonian during these processes is smaller than PA/T, and thus the adiabaticity of the interaction will not be violated and the approximate treatment given below is valid.

[4]For the derivation for the case $[P, H_D] \neq 0$ see Dass and Qureshi (1999).

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[5].

This strict mathematical result can also be understood in terms of the adiabatic theorem and the first order perturbation theory in quantum mechanics. By the adiabatic theorem, the adiabatic interaction during the protective 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: $\delta E = \langle H_I \rangle = P\langle A\rangle_n/T$, and the corresponding time evolution $e^{-iP\langle A\rangle_n/\hbar}$ then shifts the pointer by the expectation value $\langle A\rangle_n$.

3 Physical implications of protective measurement

What are the physical implications of protective measurement?[6] An immediate implication is that the result of a protective measurement, namely the expectation value of the measured observable in the measured state, reflects the actual physical property of the measured system, as the system is not disturbed after this result has been obtained[7]. This is in accordance with the fundamental assumption that the result of a measurement that does not disturb the measured system reflects the actual property or state of the system. Moreover, since the wave function can be reconstructed from the expectation values of a sufficient number of observables, the wave function of a quantum system is a representation of the physical state (or ontic state) of the system[8]. This is a further

[5]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), though these objections have been answered (Aharonov, Anandan and Vaidman 1996; Dass and Qureshi 1999; Vaidman 2009; Gao 2012).

[6]Several authors, including the inventors of protective measurements, have given some initial analyses of the implications of protective measurement (Aharonov 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 single system. In particular, they thought that the wave function describes 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.

[7]For a realistic protective measurement 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, and after obtaining the result the measured state also collapses to the state $\perp$. However, the key point here is that when the measurement obtains the expectation value of the measured observable, the state of the measured system is not disturbed. Moreover, the above probability can be made arbitrarily small in principle when T approaches infinity, as well as negligibly small in practice by making T sufficiently large.

[8]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

implication of protective measurements[9].

Let's try to find what physical state the wave function represents. According to quantum mechanics, a quantum system being in a position eigenstate has a definite position in space[10]. Moreover, since the system has properties such as mass and charge, the mass and charge of the system also exist in the definite position. Then the mass and charge of a system concentrating on a definite position can be regarded as the physical state of the system represented by one of its position eigenstates. Now that the wave function of a quantum system is a representation of the physical state of the system, the mass and charge of a quantum system being in a position superposition state should be distributed throughout all positions in superposition, and the physical state of the system will be the mass and charge distributions in space. The existence of such mass and charge distributions can be seen more clearly from the interaction terms in the Schrödinger equation. 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 does not distribute in some regions where the wave function is nonzero, then there will not exist electrostatic interaction there[11].

In the following, we will show that the existence of mass and charge distributions for a quantum system is further supported by an analysis of measurements (which do not disturb the measured system), and in particular, protective measurements can actually measure the mass and charge distributions of a quantum system in space. No matter how to define measurement, a measurement must be realized by certain interaction between the measured system and the measuring device. Concretely speaking, the measuring device is influenced by the measured system through an interaction that depends on the measured property, and the change of the measuring system then reflects the measured property of the measured system. For example, a position measurement must depend on

physical state, though the wave function 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 projective 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.

[9]Note that this implication is independent of whether the wave function of the system is known beforehand for protective measurements. For even though we know the wave function, which is an abstract mathematical object, we still don't know its physical meaning.

[10]In standard quantum mechanics, the only place to tell what properties a single system possesses independent of observation is the so-called the eigenvalue-eigenstate link. It says that an observable pertaining to a given system has a value if and only if the system is in the corresponding eigenstate of that observable. Although in a realistic alternative to quantum mechanics, an observable may also have a value even if the system is not in the corresponding eigenstate of that observable, no one would deny that an observable, if it exists, has a value if the system is in the corresponding eigenstate of that observable.

[11]Note that the charge distribution is not necessarily classical, and its interaction with other charges is not necessarily classical either (it is probably such a classical prejudice that prevents people from admitting the existence of the charge distribution for a charged quantum system such as an electron). In other words, its existence can be compatible with the Schrödinger equation, which describes its interaction with other charges by the interaction potential terms. We will analyze the physical nature of the charge distribution later.

the existence of certain position-dependent interaction between the system and the device such as electrostatic interaction between two electric charges. The existence of an electrostatic interaction between a charged measuring device and a measured system then tells us that the measured system has electric charge. Moreover, since the strength of the interaction relates to the distance between the two interacting systems, the measurement result may also reflect the charge distribution of the measured system in space. Then, for a projective position measurement of a position eigenstate of an electron, if the measurement interaction is electrostatic interaction between the electron and a charged measuring device, then the result of the measurement (which does not disturb the measured electron) will indicate that the electron, which is in a position eigenstate, has charge e in its position. Similarly, when a protective measurement (which does not disturb the measured system) is realized by electromagnetic or gravitational interaction between the measured system and the measuring device, the measurement can also measure the charge or mass distribution of the system, which may be in a general position superposition state. Let's give a detailed explanation.

Consider a protective measurement of the charge of a quantum system with charge Q in a small spatial region V having volume v. This is equivalent to measuring the following observable:

$$A = \begin{cases} Q, & \text{if } x \in V, \\ 0, & \text{if } x \notin V. \end{cases} \tag{14}$$

A protective measurement of A in a general superposition state $\psi(x,t)$ yields

$$\langle A \rangle = Q \int_V |\psi(x,t)|^2 dv, \tag{15}$$

which gives the charge of the system in the region V. When $v \to 0$ and after performing measurements in sufficiently many regions V, we can find the charge density everywhere in space, which turns out to be $\rho_Q(x,t) = Q|\psi(x,t)|^2$[12].

This result can be illustrated by a specific example. Consider a quantum system with charge Q whose wave function is

$$\psi(x,t) = a\psi_1(x,t) + b\psi_2(x,t), \tag{16}$$

where $\psi_1(x,t)$ and $\psi_2(x,t)$ are two normalized wave functions respectively localized in their ground states in two small boxes 1 and 2, and $|a|^2 + |b|^2 = 1$. A measuring electron, whose initial state is a Gaussian wave packet narrow in both position and momentum, is shot along a straight line near box 1 and perpendicular to the line of separation between the boxes. The electron is detected on a screen after passing by box 1. Suppose the separation between the boxes is large enough so that a charge Q in box 2 has no observable influence on the electron. Then if the system is in box 2, namely $|a|^2 = 0$, the trajectory of the electron wave packet will be a straight line as indicated by position "0" in Fig.1, indicating that there is no charge in box 1. If the system is in box 1, namely $|a|^2 = 1$, the trajectory of the electron wave packet will be deviated by the

[12]Similarly, we can protectively measure another observable $B = \frac{1}{2i}(A\nabla + \nabla A)$. The measurements will give the electric flux density $j_Q(x,t) = \frac{Q}{2i}(\psi^*\nabla\psi - \psi\nabla\psi^*)$ everywhere in space.

electric field of the system by a maximum amount as indicated by position "1" in Fig.1, indicating that there is a charge Q in box 1. These two measurements are conventional measurements of the eigenstates of the system's charge in box 1, and their results can reveal the actual charge distribution in box 1. However, when $0 < |a|^2 < 1$, i.e. when the measured system is in a superposition of two eigenstates of its charge in box 1, it is well known that such conventional measurements cannot detect the actual charge distribution in box 1.

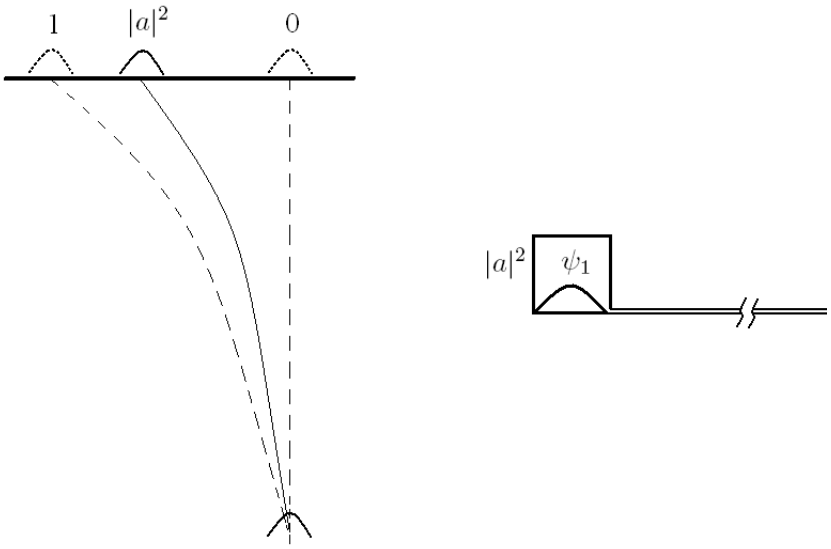

Fig.1 Scheme of a protective measurement of the charge distribution of a quantum system

Now let's make a protective measurement of the charge of the system in box 1 for the general superposition state. Since the state $\psi(x,t)$ is degenerate with its orthogonal state $\psi^{'}(x,t) = b^*\psi_1(x,t) - a^*\psi_2(x,t)$, we need an artificial protection procedure to remove the degeneracy, e.g. joining the two boxes with a long tube whose diameter is small compared to the size of the box[13]. By this protection $\psi(x,t)$ will be a nondegenerate energy eigenstate. The adiabaticity condition and the weakly interacting condition, which are required for a protective measurement, can be further satisfied when assuming that (1) the measuring time of the electron is long compared to $\hbar/\Delta E$, where ΔE is the smallest of the energy differences between $\psi(x,t)$ and the other energy eigenstates, and (2) at all times the potential energy of interaction between the electron and the system is small compared to ΔE. Then the measurement by means of the electron trajectory is a protective measurement, and the trajectory of the electron wave packet is only influenced by the expectation value of the charge of the system in box 1. As a result, the electron wave packet will reach the position "$|a|^2$" between "0" and "1" on the screen as denoted in Fig.1, indicating that there is a charge $|a|^2Q$ in box 1.

[13]It is worth stressing that the added protection procedure depends on the measured state, and different states need different protection procedures in general. This means that a protective measurement with an artificial protection procedure requires that the wave function of the measured system is known beforehand.

4 Meaning of the wave function

We have argued that according to protective measurements, the mass and charge of a quantum system are distributed throughout space, and the mass and charge density in each position is proportional to the modulus squared of the wave function of the system there. In this section, we will further investigate the physical origin of the mass and charge distributions. As we will see, the answer may provide an important clue to the physical meaning of the wave function[14].

4.1 The mass and charge distributions are effective

There are two good motivations for our further investigation. The first one is that although the existence of mass and charge distributions can be extended to a many-body system as Schrödinger (1926) originally shown, the distributions contain no information about the entanglement between the sub-systems of the many-body system. This indicates that the mass and charge distributions have a deeper physical origin, and a further analysis of the origin is needed in order to know exactly what physical state the wave function represents. The second motivation is that the mass and charge distributions have two possible forms (as will be shown below), while there is only one actual form of the distributions; we need to determine which possible form is the actual one.

As argued earlier, protective measurements show that the expectation values of observables are the properties of a quantum system. These properties are defined either at a precise instant or during an infinitesimal time interval. Correspondingly, the mass and charge distribution of a quantum system, which can be protectively measured as the expectation values of certain observables, has two possible existent forms: it is either real or effective. The distribution is real means that it exists throughout space at the same time. The distribution is effective means that at every instant there is only a localized, point-like particle with the total mass and charge of the system, and its motion during an infinitesimal time interval forms the effective distribution. Concretely speaking, at a particular instant the mass and 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 mass and charge density. Moreover, the motion of the particle is ergodic in the sense that the integral of the formed mass and charge density in any region is required to be equal to the expectation value of the total mass and charge in the region.

In the following, we will determine the existent form of the mass and charge distribution of a quantum system. If the mass and charge distribution is real, then any two parts of the distribution (e.g. the two wavepackets in box 1

[14]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 (cf. Ney and Albert 2013). 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.

and box 2 in the example given in the last section), like two electrons, will have gravitational and electrostatic interactions described by the interaction potential terms in the Schrödinger equation[15]. The existence of such gravitational and electrostatic self-interactions for individual quantum systems 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 charge distribution of an electron is incompatible 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 the mass and charge distribution is effective, there will exist no gravitational and electrostatic self-interactions of the effective distribution, as there is only a localized particle at every instant. This is consistent with the superposition principle of quantum mechanics and the Schrödinger equation.

A further clarification may be needed in order to understand the above analysis. It can be seen that the non-existence of self-interaction of the mass and 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 in the example given in the last section, the two charges in box 1 and box 2 have no such electrostatic interaction. This puzzle is not so much dependent on the existence of mass and charge distributions as properties of a quantum system. It is essentially that according to quantum mechanics, the wavepacket ψ_1 in box 1 has interaction with any test electron (e.g. deviating the trajectory of the electron wavepacket), so does the wavepacket ψ_2 in box 2, but these two wavepackets, unlike two electrons, have no interaction.

Facing this puzzle one may have two choices. The first one is simply admitting that the non-existence of self-interaction of the mass and 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 mass and charge distribution[16]. The mass and charge distribution has two possible 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 mass and charge distribution may also help explain the non-existence of self-interaction of the distribution.

[15]Moreover, these two parts will be also entangled and their wave function be defined in a six-dimensional configuration space.

[16]An immediate explanation may be that why the two wavepackets with charges have no electrostatic interaction is because they belong to one quantum system such as an electron, and if they belong to two charged quantum systems such as two electrons, then they will have electrostatic interaction. However, this explanation seems still unsatisfactory, and one may further ask why two wavepackets of a charged quantum system such as an electron, each of which has charge, have no electrostatic interaction.

This is just what the previous analysis has done. The analysis not only explains the non-existence of self-interaction of the mass and charge distribution but also determines the existent form of the distribution. The reason why two wavepackets of an electron, each of which has part of the electron's charge, have no electrostatic interaction is that these two wavepackets do not exist at the same time, and their 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, there exists no electrostatic self-interaction of the 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[17].

To sum up, we have argued that the superposition principle of quantum mechanics requires that the mass and charge distribution of a quantum system such as an electron is not real but effective; 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, and the mass and charge density in each position is proportional to the modulus squared of the wave function of the system there.

4.2 The ergodic motion of a particle is discontinuous and random

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 the effective mass and charge distribution during a finite time interval. But according to quantum mechanics, the effective mass and charge distribution is 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 boxes $\psi_1(x) + \psi_2(x)$. 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 $e|\psi_1(x) + \psi_2(x)|^2$[18].

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)[19]. This is understandable,

[17]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 describes the gravitational self-interaction of classical mass density.

[18]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.

[19]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.

since that the mass and charge density is large in a position requires that 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[20].

In summary, 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 square of its wave function there. In other words, the ergodic motion of the particle is random and discontinuous.

4.3 Interpreting the wave function

According to the above analysis, microscopic particles such as electrons 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 2011). 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[21]. This is similar to the description of a classical fluid in hydrodynamics. But their physical meanings are different. The particle cloud is formed by the random discontinuous motion of a single particle, and the density of the cloud, $\rho(x,t)$, represents the probability density for the particle to appear in position x at instant t. By assuming that the nonrelativistic equation of motion is the Schrödinger equation in quantum mechanics, the complex wave function $\psi(x,t)$ can be uniquely expressed

[20]Note that the propensity here denotes single case propensity. For 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. Like the Copenhagen interpretation of quantum mechanics, these interpretations cannot be wholly satisfactory because of resorting to the vague concept of measurement.

[21]As noted before, 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 charge density and electric flux density of the cloud, respectively.

by $\rho(x,t)$ and $j(x,t)$ (except for an overall phase factor):

$$\psi(x,t) = \sqrt{\rho(x,t)}e^{im\int_{-\infty}^{x}\frac{j(x',t)}{\rho(x',t)}dx'/\hbar}. \tag{17}$$

In this way, 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 also help 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 near the instant, 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 of particle 1 appearing in position x_1 and particle 2 appearing in position x_2, ..., and particle N appearing in position x_N[22]. Since these two quantities are defined not in the real three-dimensional space, but 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[23]. In particular, the modulus square of the wave function determines the probability density of the particles appearing in certain positions in space. By contrast, the density and flux density of the particle cloud, which are defined during an infinitesimal time interval near a given instant, 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.

4.4 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[24]. 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

[22] 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 $\rho(x_1, x_2, ...x_N, t) = \prod_{i=1}^{N}\rho(x_i, t)$.

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

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

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 square of its wave function and consistent with quantum mechanics [25].

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[26]. 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.

[25]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 a 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 wavefunction collapse is needed to solve the measurement problem for a theory of random discontinuous motion of particles. For details see Gao (2011).

[26]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.

5 A paradigm shift in understanding quantum mechanics

Protective measurement is a new 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 lead to a paradigm shift in 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[27]. Therefore, the expectation values of observables should be taken as the physical properties of a single quantum system, not those of an ensemble. This is the first conceptual shift brought by protective measurement[28].

Since the wave function can be reconstructed from the expectation values of a sufficient number of observables, this shift 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. However, all these are only the beginning.

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 (even if the system is not in an eigenstate of the measured observable)[29]. When a protective measurement is realized by electromagnetic

[27]The essential difference between performing projective measurements on an ensemble of identical systems and performing a series of protective measurements on a single system is that the later may not disturb the measured system and thus can measure the actual physical state of the system. In this sense, protective measurements do gain new information about the measured system (even if the wave function of the system is known beforehand).

[28]Note that most existing ontological formulations of quantum mechanics such as the de Broglie-Bohm theory and the many-worlds interpretation have not accommodated this important result. This may be the main reason why these theories have not made sense of the wave function. It has been argued that some of these theories can hardly be consistent with this result (Gao 2011).

[29]Unfortunately most people still ignore this important message even when the concept of protective measurement has been with us for twenty years. Physical state or ontic state is still a strange concept for people who learned quantum mechanics from standard textbooks. Most physicists have been accustomed to the "fact" that there is no visualizable physical picture in quantum mechanics. Once they have been given one, no matter what it is, they may unconsciously regard it as a mere consequence of classical prejudices. Even for realists who believe in an ontological interpretation of quantum mechanics such as the de Broglie-Bohm theory, the existence of the mass and charge distribution of an electron, which is measurable by protective measurements, may be still very strange, and they may still regard it as a consequence of classical prejudices. Most of them seem to be accustomed to the situation

or gravitational interaction between the measured system and the measuring device, it can measure the charge or mass distribution of the system. It turns out that the mass and charge of a quantum system such as an electron is distributed throughout space, and the mass and charge density in each position is proportional to the modulus square of its wave function there.

The 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 requires the mass and charge distribution is effective, that is, it 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 of the particle appearing in every position is equal to the modulus square 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 in quantum mechanics 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 wavefunction collapse is part of the equation of motion[30].

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.

where the physical state of a single quantum system cannot be measured in principle and must be given a priori (the only restriction is the consistency with the probability distribution of measurement results of an ensemble of identical system). For all these people, protective measurement, if it does have the implications as discussed in this article, will no doubt induce a paradigm shift in our understandings of quantum mechanics.

[30]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 2011).

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

[12] Gao, S. (2011). Interpreting Quantum Mechanics in Terms of Random Discontinuous Motion of Particles. http://philsci-archive.pitt.edu/9057/.

[13] Gao, S. (2012). On Uffink's alternative interpretation of protective measurements. http://philsci-archive.pitt.edu/9447/.

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

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

[16] 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.

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

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

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

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

[21] 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.

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

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

[24] 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.

Questionnaire

The following are the main results of this article. If you think any result is false, please briefly explain why or give a more detailed counterargument. Thanks!

1. By a protective measurement 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 (pp.1-4).

2. The result of a protective measurement, namely the expectation value of the measured observable in the measured state, reflects the actual physical property of the measured system (pp.4).

3. Protective measurement implies that the wave function of a quantum system is a representation of the physical state (or ontic state) of the system (pp.4).

4. Protective measurements can measure the mass and charge distributions of a quantum system in space. It turns out that the mass and charge density in each position is proportional to the modulus square of the wave function of the system there (pp.5-7).

5. The charge distribution of a quantum system has two possible forms. The superposition principle of quantum mechanics requires that the charge distribution is effective, that is, it is formed by the ergodic motion of a localized particle with the total charge of the system (pp.8-9).

6. 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 of the particle appearing in every position is equal to the modulus square of its wave function there (pp.10).

7. The wave function in quantum mechanics 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 (p.11-13).