Progress of Theoretical Physics, Vol. 87, No, 2, February 1992

Theory of Measurement

—— *A Note on Conceptual Foundation of Quantum Mechanics* ——

Yutaka Toyozawa

Department of Physics, Chuo University, Tokyo 112

(Received August 28, 1991)

Theory of measurement is presented which is based on the statistical interpretation of quantum mechanics. The reduction of wave packet is ascribed to undue comparison of the state of an ensemble before measurement and the state of an individual system after measurement. By studying the evolution of statistical operators for subsystems and composite systems comprising an object, a measuring apparatus and an observer, one can definitely locate where the measurement is done, thus negating the principle of psycho-physical parallelism due to von Neumann. Time needed for measurement is also discussed. The Einstein-Podolsky-Rosen state is described as an eigenstate of correlation. The argument lends support to the standpoint that the physical state of a system is to be represented by the statistical operator, not by the wave function, in the conceptual foundation of quantum mechanics.

§ 1. Introduction

Quantum mechanics, in spite of its brilliant successes in application, has been subject to serious controversy on its conceptual foundation[1] ever since its discovery. Its paradoxical aspects have been described in various contexts.

Quantum mechanics (Q. M.) postulates that a measurement brings an object from state ψ to one of the eigenstates ψ_n of the measured observable, with probability $|(\psi_n, \psi)|^2$. This sudden and acausal transition, the so-called *reduction of wave packet*, cannot be described by time-dependent Schrödinger equation (t.d.S.e.); the latter provides only the description of causal evolution of a mechanically isolated system. Thus, von Neumann[2] postulates two kinds of process, process **1** which is *acausal* and *non-unitary*, and process **2** which is *causal* and *unitary*, both being indispensable for constructing Q.M. as a self-contained scheme of physical principles.

The statistical nature of quantum mechanical measurements, with the probability Ansatz mentioned above, has been established by all experiments performed so far, and none doubts it today. On the other hand, the reduction of wave packet by measurement was a notion so strange and unnatural to majority of physicists that it caused a great deal of controversy, including skepticism and criticism on this orthodox interpretation of Q. M. One may well think of describing the measuring process purely mechanically by solving t.d.S.e. for the composite system consisting of the object and the measuring apparatus. Von Neumann showed, however, that the result is *only to shift the undesired reduction to the next stage* where the second measuring apparatus or an observing person recognizes the recorder of the first apparatus. One can shift this reduction further on to a later stage by extending the system which is to be described causally by t.d.S.e., namely, as process **2**, but on the final stage one must invoke process **1**, the reduction, in order to decide one out of many possibilities in the linear combination. Von Neumann claimed that the boundary between "the observed" and "the observing" *can be shifted to anywhere one likes without affecting the physical consequence.* This "principle of psycho-physical parallelism" aroused the greatest discontent and disconcertion among physicists, but strange to say, there does not seem any argument plausible enough to refute it, except perhaps for the naive intuition of physicists against it.

The purpose of this paper is to present a more plausible interpretation of Q.M. which is free from the notion of the reduction of wave packet (no need of process **1**), and which refutes the psycho-physical parallelism. The author wants to put his argument, as far as possible, on the experimental facts which established the statistical nature of quantum mechanical measurement with its probability Ansatz. In this respect, the most natural standpoint is the statistical interpretation:[3] Q.M. can predict *only the behavior of a statistical ensemble, not that of an individual system,* and the wave function also represents the state of an ensemble of identical systems, not the state of an individual system. What takes place in the actual measurement is then as follows. The ensemble was in the state ψ before measurement, while measurement on one system of the ensemble results in ψ_n, and measurement on another system in ψ_m, and so on. The notorious "reduction of wave packet" is nothing but the result of *undue* comparison of the state of *the ensemble before* measurement and the state of *an individual system after* measurement. It is not surprising that comparison between concepts of different categories results in something disjunct. If one takes consistent attitude of comparing concepts of the same category, one should rather state that *measurement is a process in which the initial ensemble is decomposed into a set of subensembles each being in an eigenstate of the observable measured.* This decomposition can in fact be described by t.d.S.e.

One should now ask what is the appropriate quantity to represent or characterize the state of the ensemble. In order to find the right answer, one must look back on the history of Q. M. It was only after Schrödinger presented the wave equation that Born proposed the probability interpretation of the wave function. Namely, the notion of ensemble was introduced after the invention of the wave function. It is of interest, though somewhat ridiculous as was once denounced by a famous historian in a more general context, to imagine how it would have been if the statistical nature of quantum processes were first established and then appropriate quantity to represent the state of a statistical ensemble were looked for. People nowadays, who are well aquainted with the notion of density matrix or statistical operator, ρ, might well answer that ρ is a more appropriate quantity, at least more general than ψ, for representing the state of an ensemble. My standpoint is that ρ is not only more appropriate or more general but is *the only quantity which can represent the state.* Some of the reasons mentioned in a previous paper[4] are as follows.

First, ρ is in one-to-one correspondence with the physical state of the ensemble while ψ cannot represent all the possible states. What is meant by "the physical state of the ensemble"? Since Q.M. can predict only the expectation values of observables, the physical state of an ensemble is characterized by a set of expectation values of all observables. Since the expectation value $\langle A\rangle$ of an observable A on the given ensemble is a linear functional of A, it is necessary and sufficient for characterizing the ensemble to specify a set of expectation values of a *maximal set of linearly* *independent observables*: A_i ($i=1, 2, \cdots$). The number of such observables —— Hermitic matrices —— should be (n^2-1) if the Hilbert space concerned were n-dimensional, which is equal to the number of independent matrix elements of the Hermitic ρ (the term -1 results from excluding the unit matrix $A_0 \equiv 1$ in the former, and from the normalization condition: $\mathrm{Tr}\rho=1$ in the latter).[5] Solving the set of simultaneous equations: $\mathrm{Tr}(A_i\rho)=\langle A_i\rangle$ ($i=0, 1, 2, \cdots, n^2-1$), one can determine the matrix ρ in one-to-one correspondence to the set of the $\langle A_i\rangle$'s. In contrast, ψ can represent only a specific class of ensembles —— a pure state. A normalized vector in the n-dimensional Hilbert space has $(2n-1)$ degrees of freedom, which is smaller than (n^2-1), the degree of freedom of the possible states, so far as $n \geqq 2$. This argument can obviously be extended to $n=\infty$. Thus, it is only for the historical reason that ψ played the primary role in representing the state, while from logical viewpoint ρ should play the fundamental role. The name "Erwartungskatalog", which Schrödinger presented to the wave function, is more duly to be reserved for the statistical operator.

The second reason comes from the obvious fact that Q.M. can describe only the object which is a subsystem of the entire universe because at least the observer must stand outside the object. Any subsystem, even if it is initially in a pure state, will inevitably be brought into a mixed state as soon as it interacts with and affects other systems outside. Since any system must have interacted with its outside systems in the past, one must describe its state by ρ instead of by ψ. This point will be further clarified in the present paper through the comparison of the ψ-description and the ρ-description of measurement. In contrast to the former, the latter description enables one to negate the controversial principle of psycho-physical parallelism and to remove the unphysical notion, the reduction of wave packet.

§ 2. The description of measurement by the wave function

Let us start with recapitulating the conventional description[2] of the measuring process in terms of the wave function, for later comparison with the description by statistical operator. Denote by a_n and $\psi_n(x)$ the n-th eigenvalue and eigenfunction, respectively, of an observable A of the object —— system **I**. Assume that the system **I** is in state $\psi(x)$ before measurement, and expand ψ as

$$\psi(x)=\sum_n c_n\psi_n(x)\,, \quad c_n \equiv (\psi_n, \psi)\,. \tag{1}$$

Q.M. postulates that the measurement of A brings the system into one of the eigenstates, ψ_n, abruptly and acausally with the probability $|c_n|^2$. This change of the state is called the reduction of wave packet, and is classified by von Neumann as process **1**.

If one likes, however, one can alternatively describe this measurement as a mechanical process in which system **I** interacts with the measuring apparatus —— system **II**, the composite system (**I**+**II**) evolving according to t.d.S.e. (namely within process **2** of von Neumann). The interaction between **I** and **II** is assumed to be such that subsystem **II** initially in state $\phi(y)$ is brought, after a short time contact with subsystem **I** in state $\psi_n(x)$, into state $\phi_n(y)$ without thereby changing the state of **I**. It is also assumed that the $\phi_n(y)$'s are orthogonal to each other. This needs a certain time t_s after the contact, which will hereafter be called the separation time. Then the wave function of the composite system **I**+**II**, which initially was

$$\Psi^{(0)}(x, y)=\psi(x)\phi(y)=\sum_n c_n\psi_n(x)\phi(y) \tag{2}$$

is brought, well (by $\sim t_s$) after the contact, into

$$\Psi^{(i)}(x, y)=\sum_n c_n\psi_n(x)\phi_n(y) \tag{3}$$

due to the linearity of t.d.S.e. The superfixes (0), (i) ⋯ denote for the different stages of evolution. We will be concerned only with those evolutions which are caused by mechanical contacts between different subsystems, and will not explicitly mention the automatic evolution within each subsystem.

In the Stern-Gerlach experiment, which is the simplest example of measurement, the particles with different spin states: $\psi_n(s_z)$ ($n=\uparrow$ and $\downarrow$ represent the up and down spin states, respectively, and the variable s_z is the z-component of the spin) are differently accelerated towards the z-direction during their passage (along the x-direction) through the inhomogeneous magnetic field, so that their orbital motions will be brought from their common initial state $\phi(\boldsymbol{r})$ into differently deflected (from the x-axis) $\phi_n(\boldsymbol{r})$'s which are supposed to be spatially separated (along the z-direction) and hence orthogonal to each other at time well (by $\sim t_s$) after the passage. Thus, one can measure the observable s_z, namely, know whether $n=\uparrow$ or $\downarrow$ by finding the particle in the upper or lower orbit that is well discernible.

In the above example, system **I** (s_z) and system **II** ($\boldsymbol{r}$) represent different degrees of freedom of the same particle, both being microscopic. In typical measurement, however, we use a macroscopic apparatus for system **II** so as to amplify the microscopic difference of ψ_n into macroscopic difference of ϕ_n which is readily recognized. In short, the measuring apparatus is a mechanical device which brings its state, after its contact with the object, into macroscopically discernible states ϕ_n in one-to-one correspondence with resulting states ψ_n of the object. In order that $\psi_n(x)$'s are kept unchanged during the mechanical contact, the interaction Hamiltonian between **I** and **II** must commute with the observable A to be measured.

According to this mechanical description of measurement, the wave function of the composite system **I**+**II** after measurement stays to be a linear combination (see Eq. (3)), as it was before (see Eq. (2)). It is only when an observer outside **I** and **II** recognizes which ϕ_n of **II** is realized that the reduction of wave function (3) into one of its components takes place.

If one still dislikes this reduction, process **1**, and wants to describe this recognition process mechanically, one must take into account the observer or the second apparatus as subsystem **III** and solve t.d.S.e. for the composite system **I**+**II**+**III**. The interaction between **II** and **III** is assumed to be such that the subsystem **III** in state $\chi(z)$ is brought into $\chi_n(z)$ well after its contact with subsystem **II** in state $\phi_n(y)$, without thereby changing the state $\phi_n(y)$. Since **I** has no more contact with **II**, the wave function of the composite system **I**+**II**+**III** before the **II**-**III** contact:

$$\Phi^{(\mathrm{i})}(x,\ y,\ z)=\Psi^{(\mathrm{i})}(x,\ y)\chi(z)=\sum_n c_n\psi_n(x)\phi_n(y)\chi(z) \tag{4}$$

is brought, well after that contact, into

$$\Phi^{(\mathrm{ii})}(x,\ y,\ z)=\sum_n c_n\psi_n(x)\phi_n(y)\chi_n(z) \tag{5}$$

which is again a linear combination due to the linearity of t.d.S.e. In order to see which possibility is realized in **III**, another observer outside **I+II+III** must recognize that. This results in the reduction of wave function (5).

On the basis of this argument, von Neumann claimed the principle of psycho-physical parallelism: The boundary between the observed and the observing can be shifted to anywhere one likes (e.g., between the object and the apparatus, between the apparatus and the person who recognizes it, or even between his retina and his brain) without changing the physical consequence. However, the process **1** is absolutely necessary in the final stage in order to connect the result of measurement to the recognition. If one were to reject consistently the notion of the reduction of wave packet, one would have to extend the system indefinitely so as to finally include the entire universe, which is obviously useless since there is no more observer outside. After all, *one cannot avoid the reduction of wave placket as far as one is wedded to the description by wave function.*

§ 3. The evolution of statistical operators for subsystems

Let us now ask about the state of each subsystem at different stages (i), (ii) $\cdots$, where the composite systems are in their respective states $\Psi^{(\mathrm{i})}$, $\Phi^{(\mathrm{ii})}\cdots$. A serious difficulty inherent in the wave function description is that *no subsystem can be described without referring to other subsystems* since the former is in general correlated with the latter. However, if one is interested in a particular subsystem, say, **I**, and wants to know the expectation value of the observable R *pertaining to that subsystem only*, one obtains[6] for the state $\Psi^{(0)}$ of Eq. (2):

$$\begin{aligned}\langle R\rangle^{(0)} &\equiv \int dx\int dx'\int dy\Psi^{(0)*}(x,\ y)R(x,\ x')\Psi^{(0)}(x',\ y)\\ &=\mathrm{Tr}[\rho^{(0)}R]\,,\end{aligned} \tag{6}$$

$$\begin{aligned}\rho^{(0)}(x,\ x') &\equiv \int dy\Psi^{(0)}(x,\ y)\Psi^{(0)*}(x',\ y)\\ &=\psi(x)\psi^*(x')=\sum_n\sum_{n'} c_n c_{n'}^*\psi_n(x)\psi_{n'}^*(x')\,,\end{aligned} \tag{7}$$

and similarly for $\Psi^{(\mathrm{i})}$ of Eq. (3), $\Phi^{(\mathrm{ii})}$ of Eq. (5) and the states of further extended systems:

$$\langle R\rangle^{(\mathrm{i})}\equiv\mathrm{Tr}[\rho^{(\mathrm{i})}R] \tag{8}$$

$$=\langle R\rangle^{(\mathrm{ii})}\equiv\mathrm{Tr}[\rho^{(\mathrm{ii})}R]=\langle R\rangle^{(\mathrm{iii})}\equiv\mathrm{Tr}[\rho^{(\mathrm{iii})}R]=\cdots\,, \tag{9}$$

$$\rho^{(\mathrm{i})}(x,\ x')\equiv\int dy\Psi^{(\mathrm{i})}(x,\ y)\Psi^{(\mathrm{i})*}(x',\ y)$$

$$=\sum_n |c_n|^2 \psi_n(x)\psi_n{}^*(x')$$

$$=\rho^{(\text{ii})}(x,\ x')\equiv\int dy\int dz\Phi^{(\text{ii})}(x,\ y,\ z)\Phi^{(\text{ii})*}(x',\ y,\ z)$$

$$=\rho^{(\text{iii})}(x,\ x')=\cdots, \tag{10}$$

making use of the orthogonality of $\{\phi_n\}$ and of $\{\chi_n\}$.

The well-known implication of the above equations is as follows. The state of subsystem **I** is completely described by the statistical operator ρ (in so far as one is concerned with the observables pertaining to that subsystem only), independently from its correlations with other subsystems. For the same reason, it is useful to study also the statistical operators σ and τ of subsystems **II** and **III**, respectively, at different stages. The results are as follows:

$$\sigma^{(0)}(y,\ y')\equiv\int dx\Psi^{(0)}(x,\ y)\Psi^{(0)*}(x,\ y')=\phi(y)\phi^*(y'), \tag{11}$$

$$\sigma^{(\text{i})}(y,\ y')\equiv\int dx\Psi^{(\text{i})}(x,\ y)\Psi^{(\text{i})*}(x,\ y')$$

$$=\sum_n |c_n|^2\phi_n(y)\phi_n{}^*(y')$$

$$=\sigma^{(\text{ii})}(y,\ y')\equiv\int dx\int dz\Phi^{(\text{ii})}(x,\ y,\ z)\Phi^{(\text{ii})*}(x,\ y',\ z)$$

$$=\sigma^{(\text{iii})}=\cdots, \tag{12}$$

$$\tau^{(\text{i})}(z,\ z')\equiv\int dx\int dy\Phi^{(\text{i})}(x,\ y,\ z)\Phi^{(\text{i})*}(x,\ y,\ z')$$

$$=\chi(z)\chi^*(z'), \tag{13}$$

$$\tau^{(\text{ii})}(z,\ z')\equiv\int dx\int dy\Phi^{(\text{ii})}(x,\ y,\ z)\Phi^{(\text{ii})*}(x,\ y,\ z')$$

$$=\sum_n |c_n|^2\chi_n(z)\chi_n(z')$$

$$=\tau^{(\text{iii})}(z,\ z')=\cdots. \tag{14}$$

It will also turn out to be useful to study the statistical operator of composite system **I**+**II**, before and after its contact with another subsystem, **III**. From Eqs. (4) and (5), one immediately obtains

$$\Xi^{(\text{i})}(x,\ y;\ x',\ y')\equiv\int dz\Phi^{(\text{i})}(x,\ y,\ z)\Phi^{(\text{i})*}(x',\ y',\ z)$$

$$=\sum_n\sum_{n'} c_n c_{n'}{}^*\psi_n(x)\phi_n(y)\psi_{n'}{}^*(x')\phi_{n'}{}^*(y'), \tag{15}$$

$$\Xi^{(\text{ii})}(x,\ y;\ x',\ y')\equiv\int dz\Phi^{(\text{ii})}(x,\ y,\ z)\Phi^{(\text{ii})*}(x',\ y',\ z)$$

$$=\sum_n |c_n|^2\psi_n(x)\phi_n(y)\psi_n{}^*(x')\phi_n{}^*(y')=\Xi^{(\text{iii})}=\cdots. \tag{16}$$

It is interesting to see how the statistical operators of subsystems or composite systems change when they contact mechanically with each other. First, the statistical operators for pure states are always brought into mixtures as soon as the subsystem or the composite system concerned contacts with other systems: ρ and σ change simultaneously between the stages (0) and (i) when **I** and **II** contact with each other (compare Eq. (7) with Eq. (10), and Eq. (11) with Eq. (12)), τ changes between the stages (i) and (ii) when **III** contacts with **II** (compare Eq. (13) with Eq. (14)), and Ξ changes between the stages (i) and (ii) when **II** of (**I**+**II**) contacts with **III** (compare Eq. (15) with Eq. (16)). That is the general property of the statistical operator well-known since its first introduction by von Neumann himself.

Second, a less general feature which, however, is of vital importance for measurement is pointed out: the statistical operator, *once brought into mixture, does not change any more* even if the relevant system contacts directly or indirectly with other systems. For instance, neither ρ nor σ changes after the stage (i) by direct or indirect contact of **II** with further systems (see Eqs. (10) and (12)), and neither τ nor Ξ changes after the stage (ii) by direct or indirect contact of **III** with further systems.

The third feature which is also characteristic of measurement is that subsystems (and also composite systems) are successively brought into mixtures which have common statistical weights: $|c_n|^2$, being governed by the initial state (1) of the object.

§ 4. Physical implications

We are now in a position to discuss the physical implications of the foregoing arguments. One should first note that the "primary" contact between **I** and **II** has a physical consequence essentially different from subsequent contacts between **II** and **III**, between **III** and **IV** and so on. In fact, the coherence (existence of definite phase relationship) or interferability among different states ψ_n of the object (see Eq. (1)) is lost within time $\sim t_s$ after the primary contact, and the resulting mixture is no more affected by subsequent indirect contacts with **III** and further systems. Subsystem **II** is also brought into a mixture in the primary contact, but this mixture is affected neither by direct contact with **III** nor by subsequent indirect contacts with further systems.

The *only reasonable interpretation* of these results is as follows. The measuring process is initiated by the primary contact between the object **I** and the subsystem **II** and is essentially completed at $\sim t_s$ after that (the latter statement will be elaborated later on). Namely, the measurement takes place between the stages (0) and (i). While, the subsequent contact between **II** and **III** can be interpreted to represent the recognition process since it brings **III** into a mixture (Eq. (14)) mimicking those of **II** and **I** but affecting neither of them (see Eqs. (12) and (10)). Accordingly, subsystems **II** and **III** of the present model represent the measuring apparatus and the observer, respectively, as far as their functions are concerned. Such an interpretation is fully consistent with our physical intuition. The principle of psycho-physical parallelism is definitely negated; in fact, the boundary between "the observed" and "the observing" is *unambiguously located* between **I** and (**II**+**III**), and the *measurement certainly took place between stages* (0) *and* (i).

The above argument also lends support to the standpoint[4] that the statistical operator is the only appropriate quantity to represent the physical state. In fact, the plausibility of the above argument and its consistence with our physical intuition is entirely based on the study of the statistical operator. With the wave function alone, one would have to hover around the notorious paradoxes of the measurement without being able to refute them. For instance, it is difficult to see, from the wave functions $\Psi^{(i)}$, $\Phi^{(ii)}$, $\cdots$ of the ever extending composite systems (**I**+**II**), (**I**+**II**+**III**) $\cdots$, that subsystems **I** and **II** in fact do not change any more after the stage (i).

How about the reduction of wave packet? In the Introduction, we already ascribed it to a notional phantasm resulting from the undue comparison of the state of an ensemble before measurement and the state of an individual system after measurement, and proposed to stand on the statistical interpretation in which one is concerned consistently with the behavior of a statistical ensemble instead of an individual system. This is what was done in § 3, where the evolution from one stage to another of the statistical operators of subsystems and composite systems were studied in detail. The only change of the object **I** takes place between Eqs. (7) and (10) as the result of **I**-**II** contact, whereby the initial ensemble is decomposed into subensembles with statistical weights $|c_n|^2$, each being in an eigenstate of the observable to be measured. The corresponding decomposition takes place in **II** simultaneously (see Eqs. (11) and (12)), and thereafter in **III** (see Eqs. (13) and (14)), with the same statistical weights, exactly reproducing the statistical nature of the microscopic system onto the macroscopic scale. Such set of subensembles is essentially an ensemble of classical mechanical systems in the sense that there are no interference terms characteristic of Q. M. The only thing left is the classical procedure of deciding which subensemble the result of an individual measurement belongs to. The behavior of an individual system is out of the scope of Q.M. according to the statistical interpretation.

As mentioned above, the decomposition of ensemble takes place simultaneously in subsystems **I** and **II**. The absence of interference terms in $\sigma^{(i)}$ of subsystem **II** is due to the orthogonality of the ψ_n's of subsystem **I** with which it has interacted, in exactly the same way as the absence of interference terms in $\rho^{(i)}$ of subsystem **I** is due to the orthogonality of the ϕ_n's of subsystem **II**. The former fact is of particular interest in connection with the Schrödinger's cat (subsystem **II**).[7] A feeling of something mysterious aroused by this paradoxical story is on the possibility of interference between the dead and alive states of the cat. It is impossible, however, due to the orthogonality between the decayed and the active states of the radioisotope (subsystem **I**). What is in the linear combination is the wave function of the *composite* system (Eq. (3)):

$$\Psi^{(i)} = c_a(t)\psi_{\text{active}}\phi_{\text{alive}} + c_d(t)\psi_{\text{decayed}}\phi_{\text{dead}}\,,$$

while the cat itself is simply in the mixture of alive and dead states given by Eq. (12) and is never in the linear combination: $\phi^{(i)} = c_a(t)\phi_{\text{alive}} + c_d(t)\phi_{\text{dead}}$. The fascinating thought of the superposition of macroscopically different states[8] must be distinguished from the problem of Schrödinger's cat, although the latter was one motivation for the former.

It is instructive to compare the statistical operator of composite system **I**+**II** with those of subsystems **I** and **II**. Of particular interest is the stage (i): While **I**+**II** is in a pure state (see Eq. (15)), **I** and **II**, individually, are in the mixed states (see Eqs. (10) and (12)). The expectation values of the observables R pertaining to **I** alone and S pertaining to **II** alone can be calculated rightly either with $\Xi^{(i)}$ or with $\rho^{(i)}$ and $\sigma^{(i)}$, but those of the observables X pertaining to both **I** and **II** such as $X \equiv R \cdot S$, and hence the correlation such as $\langle (R-\langle R\rangle)\cdot(S-\langle S\rangle)\rangle$, can be calculated only with the former (an example will be given in § 5). The interference terms $(n \neq n')$of Ξ in the stage (i) are removed in the stage (ii) (see Eq. (16)) as the result of **II**-**III** contact. At first glance, this fact contradicts with the foregoing statement that the measurement is performed between (0) and (i). However, there is no contradiction. As far as one is interested in the behavior of the composite system (**I**+**II**) including the correlations between **I** and **II**, one must treat (**I**+**II**) as a single system. The "measurement on the extended object (**I**+**II**)" is certainly performed between (i) and (ii) as the result of (**I**+**II**)-**III** contact (compare Eqs. (15) and (16)), which is also consistent with the foregoing statement. The same process was called "recognition" before because our object was subsystem **I** instead of (**I**+**II**). In the case that **II** is a macroscopic system, it occurs seldom that one must consider the quantity such as $X = R \cdot S$ and hence must treat **I** +**II** as a single system.

However, the above argument has an important bearing on the recent controversies about the "irreversibility theory of the measurement".[9] This theory ascribes the reduction of a wave packet to the irreversible change of the measuring apparatus from a prepared metastable state, such as supersaturated gas in the cloud chamber, to a stable state, such as liquid droplets condensed around the ionized molecules. The interference terms of the statistical operator for the composite system consisting of the object (**I**) and the measuring apparatus (**II**) (as seen in Eq. (15)) are shown to vanish as the result of taking long time average on subsystem **II** as justified on the basis of its ergodic nature. However, this procedure of removing the interference terms was criticized because of its non-unitarity:[10] The statistical operator should always evolve in a unitary way as far as the system concerned (**I**+**II** in the present case) evolves purely mechanically with no contact with other systems. This critique is in fact irrefutable. One must admit that the interference terms of the composite system (**I**+**II**) are removed as the result of its contact with the observer (**III**). The fact that the irreversibility is not the indispensable element of measurement is obvious from the examples of Stern-Gerlach experiment on one hand and the negative result measurement[11] on the other hand.

Although the irreversibility theory cannot dissolve the fundamental difficulty of the problem of measurement, the irreversible change in the measuring apparatus plays various roles of practical importance in the measurement. The first is its function of amplifying the microscopic event to a macroscopic signal, as has been emphasized in various contexts. The contact with a microscopic system is enough to bring the metastable state of the apparatus into the stable state ——— readily recognizable macroscopic signal. Second, the irreversible change of the state of the apparatus is useful in fixing the reult of measurement without any retrogression. In this respect, the Stern-Gerlach experiment is somewhat different from other typical measurements in that the irreversible change takes place in **III** instead of in **II**: The result of measuring the spin direction can be fixed only when the particle hits the photographic plate (**III**) leaving a spot on it in an irreversible way.

In this paper, another aspect is emphasized: *Rapidity* of the irreversible change from a metastable state to a stable state; the time needed (t_R) is supposed to be significantly shorter than the time scale of thermodynamic changes among quasi-equilibrium states. While the measurement on a classical system is in general idealized to be instantaneous, the measurement on quantal system must be assumed to need a finite time since the contact between the microscopic object and the macroscopic apparatus is itself a quantal process. What is then the time needed for the measurement (t_M)? Is it the time of mechanical contact (t_c) between the object and the apparatus? The answer is that t_M is generally longer than that. In order that the object is completely brought from the pure state of Eq. (7) to the mixed state of Eq. (10) without further macroscopic change in the apparatus, the ϕ_n's must be well *orthogonal to each other* in the first place, and they must be *macroscopically stable states* with macroscopic difference in the second place. The time for measurement in the Stern-Gerlach experiment is governed by the first criterion; it is the time needed for the deflected orbits to be well separated ($t_M \sim t_s$), which is much longer than t_c, the time of passage through the inhomogeneous magnetic field. In the case of a macroscopic measuring apparatus which makes use of its irreversible process, the time needed for measurement is the time needed for the apparatus in the prepared metastable state to relax completely to the macroscopically stable state ($t_M \sim t_R$), the second criterion covering the first ($t_R \gg t_s$) due to the macroscopic nature of the apparatus. In any case, the time needed for measurement is longer than the time of **I**-**II** contact which triggered the change in **II**.

The abstract statement, as postulated in Q. M., that the measurement brings the object into one of the eigenstates of the observable instantaneously, must be corrected in two respects. In the first place, the measurement needs a finite time which depends on the measuring apparatus. Second, the state of the object may have changed from that eigenstate during this time. In the cloud chamber, for instance, one can observe the trajectory of a particle only at t_R ($\sim$time for growing up of condensed liquid droplets up to a visible size) after the rapid passage (t_c) of the particle. The author is dwelling on these stories only for the purpose of liberating Q.M. from its scholastically armed axiomatics towards a more realistic world.

§ 5. A comment on the Einstein-Podolsky-Rosen problem

Wave mechanics has been so successful that one is tempted to put an element of reality on the wave function. Many people seem to believe, consciously or unconsciously, that an individual system has its own wave function representing its state, for instance, that an individual electron has a spin of definite direction before measurement.

What Einstein, Podolsky and Rosen clarified in their "reality" argument[12] is, from my viewpoint, only the incompleteness of the wave function description, although they claimed their argument to be a proof of the incompleteness of Q.M. itself. They devised a Gedankenexperiment, in which they could conclude, without directly touching the object, that it is certainly in one state ψ and certainly in another state ψ' (e.g., the spin is directed to the z-direction on one hand and to the x-direction on the other hand). Their conclusion that this is a contradiction is entirely based on the implicit assumption that the object must have a definite wave function before measurement.

To explain their argument and our standpoint in the context of our model for measurement, let us assign our subsystems **I**, **II** and **III** to the spin $\boldsymbol{s}_1$ of the first electron, the spin $\boldsymbol{s}_2$ of the second electron and the apparatus for measuring the spin component s_{2z}, respectively. We start with the stage (i) in which (**I**+**II**) are already in the spin singlet state. Because of the invariance of this state against the rotation of coordinate axis ($xyz \to x'y'z'$), one can write its wave function as

$$\Psi^{(i)}(1, 2) = 2^{-1/2}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\} \tag{17}$$

$$= 2^{-1/2}\{\alpha'(1)\beta'(2) - \beta'(1)\alpha'(2)\}, \tag{17'}$$

where α and β with and without prime denote for the eigenstates of $s_{z'}$ and s_z, respectively. If one measures s_{2z} with the apparatus **III** under the situation that the first electron is already far away, the statistical operator for the composite system **I** +**II** is brought to a mixture (see Eq. (16)):

$$\begin{aligned}\Xi^{(ii)}(1, 2; 1', 2') = (1/2)\{&\alpha(1)\beta(2)\alpha^+(1')\beta^+(2') \\ &+ \beta(1)\alpha(2)\beta^+(1')\alpha^+(2')\}\end{aligned} \tag{18}$$

since the **II**-**III** interaction which commutes with s_{2z} (as for the condition imposed on the interaction Hamiltonian for measurement, see § 2) will keep each of $\alpha(2)$ and $\beta(2)$ unchanged. If one reorients the apparatus from z to $z' \equiv x$ direction and measures $s_{2z'} \equiv s_{2x}$, one obtains

$$\begin{aligned}\Xi^{(ii')}(1, 2; 1', 2') = (1/2)\{&\alpha'(1)\beta'(2)\alpha'^+(1')\beta'^+(2') \\ &+ \beta'(1)\alpha'(2)\beta'^+(1')\alpha'^+(2')\}.\end{aligned} \tag{19}$$

$\Xi^{(ii)}$ is an eigenstate of $s_{1z}s_{2z}$ with eigenvalue $-(1/4)$ in the sense that

$$s_{1z}s_{2z}\Xi^{(ii)} = -(1/4)\Xi^{(ii)}, \tag{20}$$

while $\Xi^{(ii')}$ is an eigenstate of $s_{1x}s_{2x}$ with eigenvalue $-1/4$. Since $\langle s_{1z}\rangle^{(ii)} = \langle s_{2z}\rangle^{(ii)} = 0$ as will be seen later, one can characterize the Einstein-Podolsky-Rosen state $\Xi^{(ii)}$ as an eigenstate of $(s_{1z} - \langle s_{1z}\rangle)\cdot(s_{2z} - \langle s_{2z}\rangle)$ whose expectation value is the correlation between the two spins; more briefly, one may call $\Xi^{(ii)}$ an *eigenstate of the correlation.*

Thus, without direct contact with the first electron, one can predict s_{1z} with certainty to be in the down or up state according as measurement of s_{2z} results in up or down state, in the first experiment. In the second experiment, one can make similar prediction on s_{1x}. However, subsystems **I** and **II**, individually, stay in the same unpolarized state before and after the **II**-**III** contact:

$$\rho^{(i)}(1, 1') = \rho^{(ii)}(1, 1') = (1/2)\{\alpha(1)\alpha^+(1') + \beta(1)\beta^+(1')\} \tag{21}$$

$$=\rho^{(\mathrm{II}')}(1,1')=(1/2)\{\alpha'(1)\alpha'^{+}(1')+\beta'(1)\beta'^{+}(1')\} \tag{21'}$$

(and similar expressions for σ).

The argument by Einstein, Podolsky and Rosen that the two experiments give results on s_1 contradictory to each other (one in the z-direction and the other in x-direction) is justified only if the following two premises are valid: [1] The first spin had a definite direction in the stage (i) (before measurement); [2] the direction of the first spin does not change by measurement on the second spin. From our standpoint, the premise [1] is not valid since the first spin was in the unpolarized state in the stage (i) (see Eq. (21)).

In spite of the fact that composite system **I**+**II** is brought into differently correlated states (compare Eqs. (18) and (19)) depending on the way of measuring **II**, subsystems **I** and **II** individually remain unchanged. Expressions (21) and (21′), derivable from different Ξ's, are mathematically equal as well as physically equivalent. There are no experiments on **I** alone nor on **II** alone by which one can distinguish them. We will come back to this problem in the next section.

§ 6. Summary and further comments

Finally, we would like to make our point clearer by summarizing what we consider is the the most appropriate interpretation of Q.M. from the conceptual viewpoint.

Q.M. describes the statistical behavior of "an ensemble of identical systems", not the behavior of "an individual system". The word "system" is to be used in the former sense unless otherwise stated. The physical state of a system is represented by the statistical operator ρ in one-to-one correspondence in the sense mentioned in § 1. ρ can be defined for any system irrespective of its correlation with its outside —— the rest of the world. The evolution of ρ of a mechanically isolated system with its own Hamiltonian is given by $\rho(t)=e^{-iHt}\rho(0)e^{iHt}$, irrespective of its correlation with the outside.

The fact that one can thus describe any system in a *self-contained* way (without worrying about its correlation with the outside), on the one hand, and its *statistical nature* revealed in the measurement, on the other hand, form *a consistent and inseparable entirety of physical principles.* This never means that the statistical nature of the relevant system originates from that of the outside system with which it was or is in interaction.

Some people seem to consider that the statistical operator

$$\rho(x,x')=\sum_n |c_n|^2\psi_n(x)\psi_n^*(x') \tag{22}$$

contains two elements of statistical nature:[13] One is inherent in Q. M., namely in each ψ_n, and the other is due to our incomplete knowledge as represented by the statistical weights: $|c_n|^2$. We do not take this dualistic interpretation. From our standpoint, ρ is the only appropriate quantity to represent the state of a system, while ψ plays a subsidiary role as a basis vector. The word "eigen*state*" should be reserved for ρ (as in Eq. (20)) while "eigen*vector*" and "eigen*function*" for ψ. For instance, any

ρ can be expanded, in terms of an arbitrary orthonormal set $\{\psi'_i(x)\}$, as

$$\rho(x, x')=\sum_{ij}\rho_{ij}\psi'_i(x)\psi'_j{}^*(x'). \tag{23}$$

The matrix $\{\rho_{ij}\}$ of this bilinear expansion can be diagonalized as $\{\rho_n\delta_{nn'}\}$ by choosing an appropriate orthonormal set $\{\psi_n(x)\}$, whence follows the form (22) with $\rho_n=|c_n|^2$. Not only the $|c_n|$'s but also the ψ_n's depend on the given state $\rho(x, x')$. Thus, the decomposition of the statistical nature into the two elements as mentioned above is not more than a mathematical artifact and has no physical meaning, from our standpoint.

For the same reason, the classification of mixtures into "proper" and "improper" mixtures[14] does not seem to us to be meaningful. Underlying these concepts is an implicit assumption that an individual system has its own wave function, which is contrary to the statistical interpretation. According to the general understanding, ρ carries information on the present state only, no information on the history how the ensemble evolved up to the present state nor on the way how the mixture was introduced (e. g., by a *projection* from a greater system upon the relevant system, or as a *statistical mixture* of different ψ_n's).[4] This can also be seen from Eqs. (21) and (21′) and the statement in the last paragraph of § 5. The statistical operator would be almost useless if it were not in one-to-one correspondence with the present state of the ensemble.

Acknowledgements

The author is grateful to Profs. Z. Maki, S. Tanaka, M. Namiki, M. Aihara and N. Tanji for useful discussions and helpful comments in various occasions.

References

1) M. Jammer, *The Philosophy of Quantum Mechanics* (J. Wiley & Sons, New York, 1974). This book provides a comprehensive critical review of controversies on the conceptual foundation of quantum mechanics.
2) J. von Neumann, *Mathematische Grundlagen der Quanten Mechanik* (Springer, Berlin, 1932).
3) D. I. Blokhintsev, *Osnovy Kvantovoi Mekhaniki* (G. T. T. L., Moscow and Leningrad, 1949). See also Chapter 10 of Ref. 1) for a comprehensive survey of the statistical interpretation.
4) Y. Toyozawa, J. Phys. Soc. Jpn. **55** (1986), 2572.
5) U. Fano, Rev. Mod. Phys. **29** (1957), 74.
6) L. Landau, Z. Phys. **45** (1927), 430.
7) E. Schrödinger, Proc. Cambridge Phil. Soc. **31** (1935), 555; **32** (1936), 446.
8) A. J. Leggett, in *Foundation of Quantum Mechanics* (Phys. Soc. Jpn., 1986), p. 287.
9) A. Daneri, A. Loinger and G. M. Prosperi, Nucl. Phys. **33** (1962), 297.
10) W. H. Furry, in *Lectures in Theoretical Physics*, ed. W. E. Britten (Univ. of Colorado Press, Boulder 1966), p. 287
11) J. M. Jauch, E. P. Wigner and M. M. Yanase. Nuovo Cim. **48B** (1967), 144.
12) A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. **47** (1935), 777.
13) R. Tolman, *The Principles of Statistical Mechanics* (Oxford Univ. Press, 1938).
14) B. d'Espagnat, *Conceptual Foundation of Quantum Mechanics*, 2nd Ed. (Adison-Wesley, 1976).