SLAC-PUB-2797
September 1981
(T/E)

NON-REDUCTION OF THE STATE VECTOR IN QUANTUM MEASUREMENT*

J. R. Fox†
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

ABSTRACT

We describe a formalism which allows us to regard the quantum-mechanical state vector as developing solely according to the Schrödinger equation and not being subject to any reduction or collapse upon measurement. On the other hand our relationship, as observers, to the state is revised upon our learning the results of measurements. By introducing joint probability operators, whose expectation values in the unreduced state yield the probabilities of obtaining specified measurement results sequentially, we are able to reproduce the predictions of orthodox quantum theory for all testable quantities.

Submitted to American Journal of Physics

---

*Work supported by the Department of Energy under contract number DE-AC03-76SF00515.

†On leave from Department of Physics, University of Tasmania, Hobart, Tasmania, Australia.

1. Introduction

According to standard quantum measurement theory,[1] if, on a system in a quantum mechanical state $|\psi\rangle$, a measurement of an observable A is made which yields the value a_i, the state is reduced to an eigenstate $|a_i\rangle$ which is the normalized projection of $|\psi\rangle$ onto the subspace of eigenstates corresponding to the eigenvalue a_i.[2] This reduction may be held to be either a physical, objectively real, process or a mental action on the part of the observer who is updating his knowledge of the system by means of the measurement and consequently modifying his description of it. For reasons outlined below, we find neither alternative particularly satisfying.

Apart from the non-physical (or physical but unexplained) nature of the reduction process, there is another unpleasing aspect. In the language of the Statistical Interpretation of quantum mechanics,[3,4] which holds that $|\psi\rangle$ describes a conceptual ensemble of similarly prepared systems, the reduction from $|\psi\rangle$ to $|a_i\rangle$ might be thought of as filtering out, from the ensemble of systems described by $|\psi\rangle$, the sub-ensemble carrying the value a_i for A, and described by $|a_i\rangle$.[4,5] By implication, there are other sub-ensembles corresponding to the other eigenvalues of A. However, this view is not tenable, for it is well known that a pure state ensemble contains no statistically distinct sub-ensembles corresponding to pure or mixed quantum states. Consequently the measurement cannot be thought of as simply filtering out those members of the state ensemble corresponding to the eigenvalue a_i; a new ensemble must be formed.[6]

These unsatisfactory features of the reduction process have stimulated us to develop an alternative viewpoint which does not require the reduction of the state but which leads to quantum-mechanical predictions identical with those of the orthodox analysis.[8] We choose to think of the state vector as describing, perhaps incompletely, an objective reality rather than a state of knowledge about a system. The objective reality is revealed to us through measurement of observables. We like to think that the choice of observables to be measured is freely ours, whereas the state of the system develops according to a well-defined equation of motion, the Schrödinger equation, at all times. We find it more satisfying to think of the state in this way, with the discontinuity being not in the state but in our predictions about future measurement outcomes upon our gaining new information from the current measurement.

Our attitude may be illustrated by a simple analogy. Assume that we are called upon to perform repetitively a task involving some skill. We may initially predict that it will take ten minutes, but find that it actually takes us twenty the first time. We then predict that it will take seventeen minutes the second time, but find it takes eighteen. Our prediction is then revised to sixteen minutes for the third time. Our predictions at each stage are modified by what we have learned; we would hardly suggest that the task has objectively become easier, that as a consequence of our exertions it has become any easier for some other person to perform it. So, in the same way, we hold that it is not the state of the system that changes in a measurement (apart from the change described by the Schrödinger equation), but the observer's relation to it.

The testable predictions of quantum mechanics all involve statements of probability or linear combinations of probabilities. Each such probability is expressed in the theory as the expectation value in the relevant state of some projection operator, which could therefore be described as a probability operator. Thus the probability that in a state $|\psi\rangle$ measurement of A will yield the value a_i is

$$Pr(a_i) = \langle\psi|\Pi(a_i)|\psi\rangle$$

where the probability operator $\Pi(a_i)$ is simply the projection operator $P(a_i)$.

Our approach involves defining a joint probability operator $\Pi(f_m,\ldots,b_j,a_i)$ whose expectation value in the state $|\psi\rangle$ (always unreduced) is the joint probability that the results $a_i,b_j,\ldots,f_m$ will be obtained in a sequence of measurements on a system in the state $|\psi\rangle$, starting with a measurement of A and ending with a measurement of F. In Section 2 we define the joint probability operators and show how to calculate, in unreduced states, conditional probabilities which agree with probabilities calculated in the orthodox picture. For simplicity the time-dependence of the quantities involved is not discussed until Section 3. In Section 4 we examine a question which appears to lead to difficulties with the formalism: what differences, if any, are there between the predictions for the future of three observers, one of whom learns the result of a certain measurement, another of whom learns of the measurement but not the result, and the last of whom does not learn of the measurement at all? The difficulty is satisfactorily overcome by appeal to a basic property of the measurement interaction. In Section 5 we note that the formalism is not suitable for calculating, following a measurement, non-testable quantities such as amplitudes. However, it was not set up for this purpose; rather, it is a supplement to the mathematical structure of vectors and operators and relations between them, designed to elucidate and demystify the relationship of state and observer. A final section emphasises this point and shows how state reduction may be regarded as being simply a notational and computational convenience, without physical significance.

2. Probability Operators and Predictions

Consider a system in a state $|\psi\rangle$ on which a measurement of an observable A is made, yielding a value a_i, after which a measurement of B is made. The orthodox view is that the measurement of A reduces or collapses $|\psi\rangle$ to $|a_i\rangle$ and that the expectation value of B in the second measurement is

$$\bar{B} = \langle a_i|B|a_i\rangle \quad .$$

Similarly, the probability of the outcome b_j in the second measurement is

$$Pr(b_j) = \langle a_i|P(b_j)|a_i\rangle$$

where $P(b_j)$ is the projector onto the subspace of the Hilbert space corresponding to the eigenvalue b_j.

Our view is that the probability of the outcome b_j for the second measurement, after a_i has been found in the first, should be predicted as the conditional probability $Pr(b_j|a_i)$ expressed as the joint probability $Pr(b_j,a_i)$ of finding a_i in the first measurement and b_j in the second, divided by the probability $Pr(a_i)$ of finding a_i in the first, these probabilities being calculated as expectation values in the state $|\psi\rangle$ of suitably chosen probability operators $\Pi(b_j,a_i)$ and $\Pi(a_i)$.

Clearly the probability that the first measurement yields the value a_i is

$$Pr(a_i) = \langle\psi|P(a_i)|\psi\rangle \quad ;$$

leading to the identification

$$\Pi(a_i) \equiv P(a_i) \quad .$$

The joint probability $Pr(b_j,a_i)$ is the expectation value in the state $|\psi\rangle$ of the operator $\Pi(b_j,a_i)$ defined as

$$\Pi(b_j,a_i) \equiv P(a_i)\ P(b_j)\ P(a_i) \quad .$$

This operator is hermitian and reduces to the more obvious form $P(b_j)\ P(a_i)$ if $P(b_j)$ and $P(a_i)$ commute. The conditional probability $Pr(b_j|a_i)$ is then given by

$$Pr(b_j|a_i) = \frac{Pr(b_j,a_i)}{Pr(a_i)} = \frac{\langle\psi|\Pi(b_j,a_i)|\psi\rangle}{\langle\psi|\Pi(a_i)|\psi\rangle} = \langle a_i|P(b_j)|a_i\rangle \quad ,$$

the orthodox result for the probability of finding the value b_j for B after a measurement of A yielding the value a_i has reduced $|\psi\rangle$ to $|a_i\rangle$. In particular, in two successive measurements of A, $Pr(a_i|a_i) = 1$ as expected (neglecting the time-dependence of the state $|\psi\rangle$). By an obvious extension, the expectation value of B conditional upon a prior measurement of A yielding the value a_i is

$$\bar{B}\big|_{a_i} = \frac{\langle\psi|B(a_i)|\psi\rangle}{\langle\psi|P(a_i)|\psi\rangle} = \langle a_i|B|a_i\rangle$$

where

$$B(a_i) \equiv P(a_i)\,B\,P(a_i) = \sum_j b_j\,\Pi(b_j,a_i) \quad .$$

It is straightforward to extend this idea to a sequence of measurements, in the state $|\psi\rangle$, of $A,B,\ldots,E$ yielding values $a_i,b_j,\ldots,e_\ell$, respectively, followed by a measurement of F. The probability of obtaining the value f_m, following such a sequence, is the conditional probability

$$\Pr(f_m|e_\ell,\ldots,b_j,a_i) = \frac{\Pr(f_m,e_\ell,\ldots,b_j,a_i)}{\Pr(e_\ell,\ldots,b_j,a_i)} = \frac{\langle\psi|\Pi(f_m,e_\ell,\ldots,b_j,a_i)|\psi\rangle}{\langle\psi|\Pi(e_\ell,\ldots,b_j,a_i)|\psi\rangle}$$

$$= \langle e_\ell|P(f_m)|e_\ell\rangle$$

where

$$\Pi(f_m,e_\ell,\ldots,b_j,a_i) \equiv P(a_i)\,P(b_j)\ldots P(e_\ell)\,P(f_m)\,P(e_\ell)\ldots P(b_j)\,P(a_i) \quad .$$

If both expectation values in the quotient are zero, the probability is zero, as this can only occur when the sequence of results $a_i,b_j,\ldots,e_\ell$ is not possible in the state $|\psi\rangle$ (or, perhaps, in any state). Again, we obtain the orthodox result.

Other testable predictions of quantum mechanics are of transition probabilities. Consider the transition probability from the state $|\psi\rangle$ to the state $|\phi\rangle$; this can be expressed as the expectation value in the state $|\psi\rangle$ of the projector onto the state $|\phi\rangle$:

$$|\langle\phi|\psi\rangle|^2 = \langle\psi|\phi\rangle\langle\phi|\psi\rangle = \langle\psi|P(\phi)|\psi\rangle \quad .$$

Following a measurement in the state $|\psi\rangle$ of A, yielding the value a_i, the transition probability would be calculated by us to be

$$\frac{\langle\psi|P(a_i)\,P(\phi)\,P(a_i)|\psi\rangle}{\langle\psi|P(a_i)|\psi\rangle} = \langle a_i|\phi\rangle\langle\phi|a_i\rangle$$

which again agrees with the orthodox result using the reduced state vector $|a_i\rangle$.

We see, then, that using unreduced states and joint probability operators we can reproduce the predictions of quantum theory in its usual form.

3. Time-dependence

Although in the discussion of the last section we assumed a time-ordering of the measurements of A,B,..., we should properly be more careful for, if the state is developing in time, the actual moment at which each measurement occurs can be of importance. The time-development of the state is affected by interactions, in particular by those with the measuring apparatus(es).

Let the A,B,...F measurements occur at (increasing) times $t_a, t_b, \ldots t_f$. Then

$$Pr(f_m, t_f | e_\ell, t_e; \ldots; a_i, t_a) = \frac{\langle\psi(t)|\Pi(f_m, t_f; e_\ell, t_e; \ldots.; a_i, t_a)|\psi(t)\rangle}{\langle\psi(t)|\Pi(e_\ell, t_e; \ldots.; a_i, t_a)|\psi(t)\rangle}$$

with

$$\Pi(f_m, t_f; e_\ell, t_e; \ldots.; a_i, t_a) = P(a_i, t_a) \ldots. P(e_\ell, t_e)\, P(f_m, t_f)\, P(e_\ell, t_e) \ldots. P(a_i, t_a)$$

and

$$P(a_i, t_a) = U(t, t_a)\, P(a_i)\, U^{-1}(t, t_a)$$

where $U(t,t_a)$ is the unitary time-development operator satisfying the Schrödinger equation. This form will be found to agree with the orthodox result once again. In particular, $Pr(a_i,t_{a'}|a_i,t_a)$ is close to unity, as expected, if the change in $|\psi\rangle$ between times t_a and $t_{a'}$ is small. The implicit dependence on t of the Schrödinger picture operators $P(a_i,t_a)$ is a slightly unpleasant feature, as is the uncertainty as to what time t should be chosen at which to evaluate the expectation values (in fact, the choice is arbitrary).

The expressions, as might be expected, look more natural when written in terms of Heisenberg operators

$$\Pi_H(f_m,t_f;\ldots;a_i,t_a) \equiv P_H(a_i,t_a) \ldots P_H(f_m,t_f) \ldots P_H(a_i,t_a) \quad ,$$

where

$$P_H(a_i,t_a) = U^{-1}(t_a,0)\,P(a_i)\,U(t_a,0) \quad ,$$

and the Heisenberg state

$$|\psi\rangle_H \equiv |\psi(0)\rangle_S \quad .$$

However, despite the naturalness of the Heisenberg picture, we shall use the more familiar Schrödinger picture in the next section when discussing the interaction of the system with the measuring apparatus.

4. Possible Difficulties

Consider our original two-measurement sequence: measurement of A then of B. Let there be three observers, O_1, O_2, and O_3; the first learns that the first measurement yields the value a_i, the second learns that a measurement of A has been made but not the result, and the third is not informed of the measurement at all. What are their predictions for the probability that the measurement of B will yield the value b_j? In analysing this question we ignore the time arguments t_a, t_b and also assume for simplicity that the A-measurement is able to distinguish between all possible values of A (if this assumption is not made the argument can be appropriately adapted).

The first observer O_1 calculates the probability of finding b_j as

$$Pr_1(b_j) = Pr(b_j|a_i) = \frac{\langle\psi|\Pi(b_j,a_i)|\psi\rangle}{\langle\psi|\Pi(a_i)|\psi\rangle} = \langle a_i|P(b_j)|a_i\rangle$$

This probability must be interpreted as the relative frequency with which the result b_j is obtained in a series of experiments in which the first measurement yields the result a_i.

O_2 calculates the probability of finding b_j in the second measurement as the sum over joint probabilities for each possible outcome of the A-measurement

$$Pr_2(b_j) = \sum_i Pr(b_j,a_i) = \sum_i \langle\psi|\Pi(b_j,a_i)|\psi\rangle$$

$$= \sum_i \langle\psi|P(a_i)|\psi\rangle\langle a_i|P(b_j)|a_i\rangle$$

Naturally this probability is different from that calculated by O_1; it is the relative frequency with which the result b_j will be obtained in a series of experiments in which all values of A are accepted in the A-measurement. The difference between the two predictions is readily understandable and causes no problem.

The apparent difficulty arises when we consider the prediction of O_3:

$$Pr_3(b_j) = \langle\psi|\Pi(b_j)|\psi\rangle = \langle\psi|P(b_j)|\psi\rangle$$

which, in general, is different from the correct prediction of O_2. Why should knowledge of the measurement without knowledge of the result give O_2 an advantage over O_3?

Actually we have been unfair to O_3. For him to have a chance of making a correct prediction he needs to know how the (unreduced) state $|\psi\rangle$ has developed in time up to the B-measurement. Therefore he needs to know the details of the interaction with the A-measuring apparatus. But this is still not good enough, for, calculating with the given interaction Hamiltonian, he will find a pure state of the system, $\sum_i c_i |a_i\rangle$ say, after the A-measurement and will use this to calculate $Pr_3(b_j)$. However, for the measuring apparatus to be capable of distinguishing between all possible values of A, the state of (system + measuring apparatus) must, after the measurement, be a vector in the product Hilbert space of the form

$$|\Psi\rangle = \sum_i c_i |a_i\rangle |m_i\rangle$$

with the $|m_i\rangle$ mutually orthogonal measuring apparatus states.[1,10] O_3 will calculate

$$Pr_3(b_j) = \sum_{i,k} c_i^* c_k \langle a_i | P(b_j) | a_k \rangle \quad ,$$

but should have (and would have, if <u>all</u> details of the measurement interaction short of the result had been revealed to him) calculated

$$Pr_3(b_j) = \sum_{i,k} c_i^* c_k \langle a_i | \langle m_i | P(b_j) \otimes \mathbb{1} | a_k \rangle | m_k \rangle$$

$$= \sum_i c_i^* c_i \langle a_i | P(b_j) | a_i \rangle = Pr_2(b_j)$$

where $\mathbb{1}$ is the unit operator in the measuring apparatus Hilbert space of states.

Therefore, provided O_3 is informed about the full state vector $|\Psi\rangle$ of (system + measuring apparatus) after the A-measurement, he will correctly predict, together with O_2,

$$Pr_{2,3}(b_j) = \langle\Psi|\Pi(b_j)|\Psi\rangle$$

where here $\Pi(b_j) \equiv \Pi(b_j)\otimes\mathbb{1}$. The prediction is not conditional upon the A-measurement, although the time development of $|\Psi\rangle$ depends on the nature of that measurement. Notice that it is immaterial whether or not the result of the measurement has been noted by other observers; an observer's prediction is conditioned (modified) only when he learns a measurement result.

Apart from measurement interactions there may be others which do not correlate mutually orthogonal apparatus states with the different eigenvalues of A. Such interactions cannot lead to a measurement of A and affect predictions only to the extent that they affect the state vectors through the Schrödinger equation.

5. Limitations of the Formalism

We note briefly that whereas our joint probability operators and unreduced states reproduce the orthodox predictions for testable quantities, this is not the case for all quantum-mechanical quantities.

For example, proceeding as before, we might suggest that the overlap of $|\psi\rangle$ with $|\phi\rangle$, following a measurement of A in the state $|\psi\rangle$ yielding the value a_i, is

$$\langle\phi|\Pi(a_i)|\psi\rangle \quad .$$

But in general this differs by a numerical factor $\langle a_i|\psi\rangle$ from the orthodox result $\langle\phi|a_i\rangle$.

There appears to be no natural way to reproduce such values in our formalism. However, this does not concern us, for we did not set out to incorporate all of quantum mechanics in it. Rather, it provides a way of calculating the testable predictions of the theory which allows us to think of the system (and any measuring apparatus) as developing strictly according to the equation of motion, and to see any discontinuous aspect of the measurement process as residing in the observer and his state of knowledge.

6. Conclusions

We have argued that it is consistent to think of the state of a system as being unreduced by the measurement process, provided predictions about future measurement results are modified in the light of knowledge gained. Probability operators were defined whose expectation values in the unreduced states could be used to compute such modified predictions, consistent with those of the orthodox treatment.

With the state of the system un-reduced in a measurement of A, say, it might be thought that a repeated measurement of A need not yield the same result as the first. However, the formalism correctly predicts a high (or unit) probability of the same value recurring, if the state vector has changed only slightly (or not at all) from the time of the first measurement to the second.

We have shown that provided the observer has full knowledge of the time-development of the system and any apparatus interacting with it, he is able to make predictions which need be <u>discontinuously</u> revised only when he learns the result of some measurement. If he does not learn that result his predictions are not discontinuously revised, but in this case they refer to experiments which do not discriminate between the results that could occur in that measurement. There should be, of course, in either case, a continuous revision of predictions due to the continuous time-development of the state both between and during measurements.

We make a final point. In a long sequence of measurements it very soon becomes notationally unwieldy to use unreduced states and joint probability operators. For convenience we can consolidate our knowledge, represented by the sequence of results $a_i, b_j, \ldots e_\ell$, by using an updated state vector

$$\frac{P(e_\ell) \ldots P(b_j)\, P(a_i)\, |\psi\rangle}{\| P(e_\ell) \ldots P(b_j)\, P(a_i)\, |\psi\rangle \|} \equiv |e_\ell\rangle$$

and joint probability operators from which the arguments $a_i, b_j, \ldots e_\ell$ have been removed. A situation in which the consolidation appears to be particularly appropriate is when, for example, our B-measuring apparatus is such that the system will not pass through it unless A has a particular value, a_i say. In such a case, conditional probabilities $\Pr(b_j | a_k)$, $k \neq i$, are not testable at all, and in the state $|\psi\rangle$ the second measurement is, strictly speaking, not one of B, but of B "localised" to a particular value of A. In thinking about this, it is helpful to take A as being a position observable.

The orthodox view of measurement involves the consolidation we have described after each step in the sequence of measurements; from our point of view, however, it is nothing more than a convenience, and is without physical significance.

Of what use is any of the foregoing? We suggest that, apart from offering a more satisfactory interpretation, already alluded to, of the relationship between the state of the system and the observer, it may be of some assistance to those seeking to construct hidden variable theories. Not only is there no sudden reduction upon measurement to be accounted for, but, because of its absence, any particular measurement process must now be characterized, and most naturally so, by the observables which may be measured by it, rather than by the states to which the initial state may be reduced by it, as is more usual.[2] We have found this characterization to be of value to us in our own thinking about such a program.

FOOTNOTES AND REFERENCES

1. J. Von Neumann, "The Mathematical Foundations of Quantum Mechanics," Princeton Univ. Press, Princeton, NJ, 1955.
2. F. J. Belinfante, "A Survey of Hidden-Variables Theories," Pergamon Press, Oxford, 1973.
3. L. E. Ballentine, Rev. Mod. Phys. 42, 358 (1970).
4. J. L. Park, Am. J. Phys. 36, 211 (1968).
5. H. Margenau, Ann. Phys. 23, 469 (1963).
6. The filtering notion may be restored to respectability if one defines a reduced density matrix for the system alone by taking the trace, in the measuring apparatus subspace, of the density matrix for the (system + apparatus) (see Refs. 5, 7). The reduced density matrix for the system describes a pure state before the

measurement, but a mixed state after it; this mixture does have sub-ensembles and can be filtered.

7. M. Cini, M. De Maria, G. Mattioli and F. Nicolo, Found. Phys. 9, 479 (1979).

8. A somewhat similar discussion embodying these features has already been given by Pearle (Ref. 9).

9. P. Pearle, Am. J. Phys. 35, 742 (1967).

10. F. Selleri and G. Tarozzi, Riv. Nuovo Cimento 4 (No. 2), 1 (1981).