Approximate measurement in quantum mechanics. I*

Mary H. Fehrs
Department of Physics, Lafayette College, Easton, Pennsylvania 18042

Abner Shimony
Laboratoire de Physique Théorique et Hautes Energies, Orsay, France†

(Received 25 May 1973)

This is the first of two papers showing that the quantum problem of measurement remains unsolved even when the initial state of the apparatus is described by a statistical operator and when the results of measurement have a small probability of being erroneous. A realistic treatment of the measurement of observables of microscopic objects (e.g., the position or the spin of an electron) by means of observables of macroscopic apparatus (e.g., the position of a spot on a photographic plate) requires the consideration of errors. The first paper considers measurement procedures of the following type: An initial eigenstate of the object observable leads to a final statistical operator of the object plus apparatus which describes a mixture of "approximate" eigenstates of the apparatus observable. It is proved that each of a large class of initial states leads to a final statistical operator which does not describe any mixture containing even one "approximate" eigenstate of the apparatus observable.

I. INTRODUCTION

Several writers[1] have tried to solve the quantum-mechanical problem of measurement through one or both of the following proposals: (a) describing the initial state of the measuring apparatus by a projection onto a subspace of the associated Hilbert space or by a statistical operator, thus taking into account the practical impossibility of knowing the exact quantum state of a macroscopic object; (b) recognizing that there may be some physical inaccuracy, such as a small error in the position of a pointer needle, in the final registration of the outcome of the measurement by the apparatus. These writers hope that in a measuring process which satisfies (a) and (b), any initial state of the object will result in a final statistical state of the object plus apparatus which is a mixture of exact or approximate eigenstates of the apparatus observable. If this hope were justified, then the quantum-mechanical problem of measurement might be resolved. The apparatus observable could be considered to have, at least with high probability, a definite though unknown value at the end of the physical process of interaction between the object and the apparatus. The consciousness of the observer would become aware of this definite value, and therefore would not have to be assigned the role of reducing a superposition.

This and the following paper on approximate measurement are a continuation of earlier work,[2-6] initiated by Wigner, strongly indicating that no satisfactory solution to the measurement problem can be obtained in the manner that has just been sketched.[7] The papers differ in their formulations of the approximate measuring procedure proposed in (b). In the present paper the following formulation is adopted: If the initial state of the object is an eigenstate of the object observable, then the final statistical state of the object plus apparatus can be described as a mixture of pure quantum states, all of which are "almost" eigenstates of the apparatus observable associated with the same eigenvalue. This is a less stringent conception of measurement than those treated in Refs. 2–6, and one might therefore conjecture that it permits any initial state of the object to eventuate in a final mixture of "almost" eigenstates of the apparatus observable. The falsity of this conjecture follows from a mathematical theorem which is proved in Sec. II and discussed in Sec. III. Even if the conjecture had been true, however, it is not clear that progress would have been made towards solving the problem of measurement; for unless each pure state of the final mixture were an exact eigenstate of the apparatus observable, a reduction of a superposition would seem to be required in order to produce an objectively definite value of this observable. For this reason, the formulation of measurement analyzed in the sequel to the present paper probably has greater philosophical interest than the one analyzed here. A further conjecture, primarily of mathematical interest, is discussed in Sec. IV.

II. PROOF OF A THEOREM

In the following discussions, $\mathcal{H}_1$ and $\mathcal{H}_2$ denote the Hilbert spaces associated with the object and the apparatus, respectively, and $\mathcal{H}_1 \otimes \mathcal{H}_2$ denotes the space associated with the composite system, the object plus apparatus. Subscripted letters E and F are used for projection operators on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and correspondingly subscripted underlined letters, $\underline{E}$ and $\underline{F}$, are used for the subspaces onto which they project. However, the projection operators onto $\mathcal{H}_1$ and $\mathcal{H}_2$ themselves are in each case denoted by 1; and the one-dimensional subspace (ray) spanned by a nonzero vector u is denoted by $\langle u\rangle$, and the associated projection operator by P_u. $\underline{M}^\perp$ is the orthogonal complement of $\underline{M}$, and $M^\perp$ is the projection operator onto $\underline{M}^\perp$. An eigenvector of a statistical operator with a nonzero eigenvalue is called a "constitutive vector." A nonzero vector u is said to be "within ϵ of being contained in the subspace $\underline{M}$" if $\|u\|^{-1}\|M^\perp u\|$ is equal to or less than ϵ; and the same designation is applied to a subspace $\underline{N}$ if every nonzero vector in $\underline{N}$ is within ϵ of being contained in $\underline{M}$.

The theorem to be proved is the following.

Hypotheses:

(i) $\{\underline{E}_m\}$ is a finite or denumerably infinite family of mutually orthogonal subspaces spanning $\mathcal{H}_1$, $\{\underline{F}_m\}$ is a family of mutually orthogonal subspaces of $\mathcal{H}_2$, and U is a unitary operator on $\mathcal{H}_1\otimes\mathcal{H}_2$;

(ii) T is a statistical operator on $\mathcal{H}_2$ such that for every m and every $v\in\underline{E}_m$, the range of $U(P_v\otimes T)U^{-1}$ is within ϵ_m of being contained in $\mathcal{H}_1\otimes\underline{F}_m$, where all ϵ_m are equal to or less than some fixed ϵ, all but N of the ϵ_m are 0, and $N\epsilon^2$ is less than 1;

(iii) $U(P_u\otimes T)U^{-1}$ has a constitutive vector within ϵ_0 of being contained in one of the subspaces $\mathcal{H}_1\otimes\underline{F}_m$, say, $m=k$.

Conclusion: u is within κ of being contained in $\underline{E}_k$, where $\kappa=(2N^{1/2}\epsilon+\epsilon_0)(1-N\epsilon^2)^{-1}$; the range of $U(P_u\otimes T)U^{-1}$ is within $\kappa+\epsilon$ of being contained in $\mathcal{H}_1\otimes\underline{F}_k$.

The proof will make use of three lemmas.

Lemma 1. Hypotheses:

(a) Same as (i) of the theorem;

(b) for each m, $U(E_m\otimes\langle\eta\rangle)$ is within ϵ_m of being contained in $\mathcal{H}_1\otimes\underline{F}_m$, where all ϵ_m are equal to or less than some fixed ϵ, all but N of the ϵ_m are 0, and $N\epsilon^2$ is less than 1;

(c) u is a vector of $\mathcal{H}_1$ such that $U(u\otimes\eta)$ is within ϵ_0 of being contained in $\mathcal{H}_1\otimes\underline{F}_k$.

Conclusion: Same as the first part of the conclusion of the theorem.

Proof: Let $u'=u/\|u\|$ and $\eta'=\eta/\|\eta\|$. Then u' can be expressed as $\sum_i c_i u_i$, where u_i is a normalized vector belonging to $\underline{E}_i$ and the sum of the $|c_i|^2$ is unity. Relabel the subspaces so that $\epsilon_1,\ldots,\epsilon_N$ are the nonzero members of $\{\epsilon_m\}$. By (c)

$$1-\epsilon_0^{\,2}\le\|(1\otimes F_k)U(u'\otimes\eta')\|^2=\left\|\sum_{i\ne k}^{N}c_i\chi_i+c_k\chi_k\right\|^2,$$

where χ_m is defined as $(1\otimes F_k)U(u_m\otimes\eta')$. By (b), $\|\chi_m\|^2$ is equal to or greater than $(1-\epsilon^2)$ for $m=k$ and equal to or less than ϵ^2 for $m\ne k$. Therefore, a lower bound h on $|c_k|$ is obtained by rewriting the foregoing inequality as

$$1-\epsilon_0^{\,2}-\left\|\sum_{i\ne k}^{N}c_i\chi_i\right\|^2-2\mathrm{Re}\left(\sum_{i\ne k}^{N}c_i\chi_i,\,c_k\chi_k\right)\le|c_k|^2\|\chi_k\|^2,$$

and then replacing the third and fourth terms on the left-hand side by their respective lower bounds, $-(1-h^2)N\epsilon^2$ and $-2(1-h^2)^{1/2}N^{1/2}\epsilon$, and $\|\chi_k\|^2$ by its upper bound 1. The condition on h (with much information thrown away in this manner) is then

$$1-\epsilon_0^{\,2}-(1-h^2)N\epsilon^2-2(1-h^2)^{1/2}N^{1/2}\epsilon\le h^2.$$

This may be rewritten as a condition on $(1-h^2)^{1/2}$, which is an upper bound on $(1-|c_k|^2)^{1/2}$. Using the resulting inequality together with the condition that $N\epsilon^2$ is less than 1, one finds that this upper bound is less than κ (the quantity defined in the conclusion of the theorem).

Lemma 2. Hypotheses:

(a) Same as (i) of the theorem;

(b) T is a statistical operator on $\mathcal{H}_2$ having range $\underline{R}$ and such that for every m and every v in $\underline{E}_m$ the range of $U(P_v\otimes T)U^{-1}$ is within ϵ_m of being contained in $\mathcal{H}_1\otimes\underline{F}_m$.

Conclusion: $U(\underline{E}_m\otimes\underline{R})$ is within ϵ_m of being contained in $\mathcal{H}_1\otimes\underline{F}_m$ for each m.

Lemma 3. If u is a nonzero vector of $\mathcal{H}_1$ and A an operator on $\mathcal{H}_2$, then every eigenvector of $P_u\otimes A$ with nonzero eigenvalue a is of the form $u\otimes\eta$, where η is an eigenvector of A with the same eigenvalue a.

Lemma 2 is a slight variation of the second lemma on p. 65 of Ref. 5, while lemma 3 is exactly the third lemma on that page.

The proof of the theorem now proceeds as follows. Let ξ be the constitutive vector of $U(P_u\otimes T)U^{-1}$ referred to in hypothesis (iii) of the theorem. Then $U^{-1}\xi$ is a constitutive vector of $P_u\otimes T$. Hence, by lemma 3, $U^{-1}\xi$ has the form $u\otimes\eta$, where η is a constitutive vector of T and therefore also a member of the range $\underline{R}$ of T. Therefore, (iii) implies that $U(u\otimes\eta)$ is within ϵ_0 of being contained in $\mathcal{H}_1\otimes\underline{F}_k$, satisfying (c) of lemma 1. We have both hypotheses of lemma 2 and hence its conclusion, which implies (b) of lemma 1. Since (a) of lemma 1 is given as hypothesis (i), we now have all the hypotheses of lemma 1 and therefore its conclusion. But this is also the first part of the conclusion of the theorem. To prove the second part of the conclusion, let σ be any nonzero vector in the range of $U(P_u\otimes T)U^{-1}$. Then for some $\rho\in\mathcal{H}_1\otimes\mathcal{H}_2$, $\sigma=U(P_u\otimes T)U^{-1}\rho$; and hence for some $\tau\in\mathcal{H}_2$, $T\tau\ne0$, $\sigma=U(u\otimes T\tau)$. We can then write σ as the sum $U(E_k u \otimes T\tau) + U(E_k^{\perp} u \otimes T\tau)$, the first term being a nonzero member of the range of $U(P_{E_k u} \otimes T)U^{-1}$. Therefore

$$\frac{\|(1\otimes F_k)^{\perp}U(u\otimes T\tau)\|}{\|U(u\otimes T\tau)\|} \leq \frac{\|(1\otimes F_k)^{\perp}U(E_k u\otimes T\tau)\|}{\|U(u\otimes T\tau)\|} + \frac{\|(1\otimes F_k)^{\perp}U(E_k^{\perp} u\otimes T\tau)\|}{\|U(u\otimes T\tau)\|}$$

$$\leq \frac{\|(1\otimes F_k)^{\perp}U(E_k u\otimes T\tau)\|}{\|U(E_k u\otimes T\tau)\|} + \frac{\|U(E_k^{\perp} u\otimes T\tau)\|}{\|u\otimes T\tau\|}$$

$$\leq \epsilon_k + \kappa ,$$

where the last step uses hypothesis (ii) and the first part of the conclusion of the theorem.

III. DISCUSSION OF THE THEOREM

A possible formulation of a procedure of measurement in quantum theory which satisfies proposals (a) and (b) of Sec. I is given by hypothesis (i) together with a somewhat modified version of hypothesis (ii), in which the condition that only N of the $\{\epsilon_m\}$ are 0 is replaced by the condition that all the ϵ_m are much less than 1. In this formulation the subspaces $\{E_m\}$ are eigenspaces associated with distinct eigenvalues of some object observable $\mathcal{O}$. The subspaces $\{F_m\}$ are eigenspaces associated with distinct eigenvalues of some apparatus observable $\mathcal{A}$. The initial state of the apparatus is described by a statistical operator T, in recognition of the fact stated in proposal (a) that the exact quantum state of the apparatus in unknown. If $u \in \mathcal{K}_1$, then $P_u \otimes T$ is a statistical operator describing the object plus apparatus at the beginning of the measurement process. $U(P_u\otimes T)U^{-1}$ is the statistical operator which evolves from $P_u \otimes T$ in a certain time interval, U being the unitary operator governing the evolution of any pure quantum state of the object plus apparatus during that interval. The modified hypothesis (ii) (with $\epsilon_m \ll 1$, but with no condition on the number of nonzero ϵ_m) asserts that the final statistical operator describes a mixture of quantum states which are all "almost" eigenstates of $\mathcal{A}$ with the same eigenvalue. Hence, if one begins with the object observable having a definite value λ_m, then the final statistical state of the object plus apparatus is such that a subsequent measurement of the apparatus observable will yield a value from which the correct original value of $\mathcal{O}$ can be inferred with a high probability. Since $\mathcal{A}$ is in practice a macroscopic observable, the measurement of $\mathcal{A}$ could presumably be performed merely by looking, or in any case without the elaborate amplifying equipment needed to measure a microscopic observable.

The theorem of Sec. II shows that a procedure of measurement which satisfies this formulation, and also the condition of hypothesis (ii) that $N\epsilon^2$ is less than 1, will not fulfill the hopes of the writers in Ref. 1. For, according to the theorem, a final statistical state of the object plus apparatus which describes a mixture of exact or approximate eigenstates of the apparatus observable will come about only if the initial state of the object is almost an eigenstate of the object observable.

The mathematical results of Refs. 2, 4, and 5, and part of Ref. 3, are contained in the theorem by letting ϵ and ϵ_0 both be 0. The result of the other part of Ref. 3 is obtained by letting ϵ be 0, but taking ϵ_0 to be nonzero but much less than 1. That of Ref. 6 is obtained by imposing the following conditions: $0 \leq \epsilon_0 \ll 1$, $0 \leq \epsilon \ll 1$, and $N=2$.

The theorem can surely be strengthened somewhat, since much information was thrown away in the course of proving it. However, the counterexample given in Sec. IV shows that the natural generalization of the theorem is false. It seems unlikely that any of the valid strengthened versions of the theorem would have any physical or philosophical implications of interest which are not already contained in the theorem as stated, unless a different formulation of the procedure of measurement is given.

Indeed, it is dubious that the results presented above have much philosophical significance, because a scheme of measurement in which the pure states of the final mixture are only "almost" eigenstates of the apparatus observable does not seem to ensure the objective existence of a definite value of the apparatus observable. A superposition of two non-null eigenvectors v_1 and v_2 of the operator A, with distinct eigenvalues, does not represent a state in which the corresponding observable $\mathcal{A}$ is definite, although possibly unknown. This is a matter of principle in the ordinary interpretation of quantum mechanics, and it is not altered when the norm of v_1 is much greater than the norm of v_2. Any attempt to dismiss as negligible the contribution from a vector of small norm must be regarded as an alteration of principle, which would require justification. Without such an alteration of principle, the scheme of measurement of this paper would not be satisfactory—even apart from the theorem of Sec. II—because it would apparently have to be supplemented by some means of reducing a superposition.

IV. FURTHER CONJECTURE

A natural generalization of the theorem of Sec. II is the following.

Conjectured theorem: Hypotheses:

(i′) Same as (i) of the theorem;

(ii′) T is a statistical operator on $\mathcal{H}_2$ such that for every m and every $v \in \underline{E}_m$, the range of $U(P_v \otimes T)U^{-1}$ is within ϵ_m of being contained in $\mathcal{H}_1 \otimes \underline{F}_m$, $\epsilon_m \ll 1$;

(iii′) $U(P_u \otimes T)U^{-1}$ has at least one constitutive vector within ϵ_0 of being contained in the subspace $\mathcal{H}_1 \otimes \underline{F}_m$, $\epsilon_0 \ll 1$.

Conclusion: u is within κ of being contained in E_k, $\kappa \ll 1$.

A counterexample shows that this conjecture is false, even when the relation $\ll$ is construed as stringently as one could reasonably demand in the hypotheses and as liberally as possible in the conclusion. Let $\{\underline{E}_m\}$, $m = 1, \ldots, N$, be a family of one-dimensional subspaces spanning $\mathcal{H}_1$, and let the unit vector u_m span E_m. Let $T = P_\eta$, where η is a unit vector of $\mathcal{H}_2$. Let $\{\chi_m, \xi\}$ be a set of $N+1$ orthonormal vectors of $\mathcal{H}_2$, and let $\underline{F}_m$ be spanned by χ_m for $m < N$, while $\underline{F}_N$ is spanned by χ_N and ξ. Finally, let v be an arbitrary unit vector of $\mathcal{H}_1$ and (partially) define the operator U by

$$U(u_i \otimes \eta) = (N+1)^{-1} v \otimes \left[N\chi_i - \sum_{j \neq i} \chi_j + (N+2)^{1/2}\xi \right].$$

U as so far defined preserves inner products, and it may be extended to a unitarity. Clearly, hypotheses (i′) and (ii′) are satisfied if N is sufficiently large, with $\epsilon_m \leq (N+1)^{-1/2}$. Now let w be defined as $N^{-1/2}\sum_1^N u_i$. Then $U(w \otimes \eta)$ is in the range of $U(P_w \otimes T)U^{-1}$, and

$$U(P_w \otimes T)U^{-1} = (N+1)^{-1} v \otimes \left[N^{-1/2} \sum \chi_i + (N^2 + 2N)^{1/2}\xi \right],$$

which is within $(N-1)^{1/2}(N^2+N)^{-1/2}$ of being contained in $\mathcal{H}_1 \otimes \underline{F}_N$. Thus, hypothesis (iii′) is satisfied, by taking N sufficiently large. But the conclusion is not satisfied, since

$$\|E^{\perp} w\| / \|w\| = (1 - N^{-1})^{1/2},$$

which is as close to 1 as one desires.

The invalidity of the conjecture opens no new avenue for a solution to the problem of measurement, for if the hypotheses (i′) and (ii′) are satisfied, one can construct initial states of the object such that the final statistical state of the object plus apparatus is far from describing a mixture of approximate eigenstates of the apparatus observable. For this purpose it suffices to choose N greater than 2 but small enough that $N\epsilon_m^2$ is much less than 1 for each $m \leq N$, and to let u be $N^{-1/2}\sum_1^N u_i$, where u_i is a unit vector belonging to $\underline{E}_i$. If one takes the $\mathcal{H}_1$ of the theorem of Sec. II to be the direct sum of the subspaces E_i, with $i = 1, \ldots, N$, and ϵ_0 to be $\frac{1}{2}$, then that theorem implies that the final statistical state of the object plus apparatus does not contain a single constitutive vector within $\frac{1}{2}$ of being an eigenstate of the apparatus observable.

ACKNOWLEDGMENT

We wish to thank Professor B. d'Espagnat for his corrections of an earlier version of this paper and for his hospitality to one of us (A.S.) at the Laboratoire de Physique Théorique et Hautes Energies at Orsay. We also wish to thank Professor A. Fine for illuminating correspondence on the conception of measurement in quantum mechanics.

---

*This paper is based upon material contained in a thesis submitted by one of the authors (M.H.F.) to the Graduate School of Boston University in 1973 in partial fulfillment of the requirements for the degree of Doctor of Philosophy. It was written while one of the authors (A.S.) was a John Simon Guggenheim Memorial Fellow.

†Permanent address: Departments of Philosophy and Physics, Boston University, Boston, Massachusetts 02215.

[1]W. Heisenberg, *Physics and Philosophy* (Harper and Row, New York, 1962), pp. 53, 54; L. Landau and E.Lifshitz, *Quantum Mechanics* (Pergamon, London, 1958), pp. 21–24.

[2]E. P. Wigner, Am. J. Phys. 31, 6 (1963). The relevant section of this essay is entitled "Critiques of the Orthodox Theory."

[3]B. d'Espagnat, Nuovo Cimento Suppl. 4, 828 (1966).

[4]J. Earman and A. Shimony, Nuovo Cimento 54B, 332 (1968).

[5]H. Stein and A. Shimony, *Foundations of Quantum Mechanics*, edited by B. d'Espagnat, (Academic, New York, 1971), p. 56 (especially pp. 64–66).

[6]B. d'Espagnat, *Conceptual Foundations of Quantum Mechanics* (Benjamin, Menlo Park, 1971), pp. 331–339.

[7]A quite different proposal for solving the measurement problem is based upon the assumption that only some of the self-adjoint operators on the Hilbert space associated with the apparatus represent physical observables. See, for example, A. Daneri, A. Loinger, and G. M. Prosperi [Nucl. Phys. 33, 297 (1962)] and K. Hepp [Helv. Phys. Acta 45, 237 (1972)]. That proposal is not discussed in this or the following paper, but is discussed in Ref. 6, pp. 278–284.