Foundations of Physics, Vol. 7, Nos. 11/12, 1977

Statistical Inference and Quantum Mechanical Measurement

Rodney W. Benoist,[1] Jean-Paul Marchand,[1] and Wolfgang Yourgrau[1]

Received April 1, 1977

We analyze the quantum mechanical measuring process from the standpoint of information theory. Statistical inference is used in order to define the most likely state of the measured system that is compatible with the readings of the measuring instrument and the a priori information about the correlations between the system and the instrument. This approach has the advantage that no reference to the time evolution of the combined system need be made. It must, however, be emphasized that the result is to be interpreted as the statistically inferred state of the original system rather than the state of the system after measurement. The phenomenon of "reduction of states" appears in this light as a consequence of incomplete information rather than the physical interaction between measured system and measuring instrument.

1. INTRODUCTION

Statistical inference is concerned with the problem of determining the most likely distribution of a probabilistic system about which partial information is available. This concept has recently been formulated in the context of quantum mechanics.(1) In the case of quantum mechanics the probabilistic system is represented by a von Neumann algebra $\mathcal{A}$, which contains the observables, and the partial information consists of (a) an a priori state v on $\mathcal{A}$, and (b) an expectation functional on a subalgebra $\mathcal{B}$ of $\mathcal{A}$.

In this paper we use this formalism in order to analyze the quantum mechanical measuring process. The significance of $\mathcal{A}$, $\mathcal{B}$, and v is as follows: Let S, A, and $S + A$ be, respectively, the systems of the measured object, the measuring apparatus, and the combination of both; then $\mathcal{A}$ contains the

[1] Departments of Mathematics and Physics, University of Denver, Denver, Colorado.

observables of $S + A$; $\mathscr{B}$, the observables of A; and v, the correlations between S and A. The formalism then permits us to determine the most likely state w of $S + A$ and, by restriction, the most likely state w_S of the measured system S.

The remarkable feature of this formulation is that the standard results of the theory of measurement (see, e.g., Refs. 2 and 3) are obtained *without making any assumptions about the time evolution of the system* $S + A$. In particular, the so-called "reduction of states" in perfect measurements appears in this light as a natural consequence of the incompleteness of our information about the system rather than its interaction with the measuring instrument.

As pointed out in Ref. 4, this view does not conflict with the fundamental assumptions of quantum mechanics. By interpreting w_S as the inferred state of S before the measurement rather than the state of S after the measurement (whatever that means), we avoid the controversy concerning unitarity and continuity of the Schrödinger time evolution.

2. STATISTICAL INFERENCE

We describe in this section the general mathematical formalism of statistical inference on which our analysis of the measuring process will be based. For the details the reader is referred to Ref. 1.

Let $\mathfrak{A}$ be a von Neumann algebra on a separable Hilbert space H, $\mathscr{B} \subseteq \mathfrak{A}$ a von Neumann subalgebra, and v a faithful normal state on $\mathfrak{A}$. Let $v_{\mathscr{B}}$ be the restriction of v to $\mathscr{B}$. We interpret the self-adjoint elements of $\mathfrak{A}$ as the observables of the system, the self-adjoint elements of $\mathscr{B}$ as the measured observables, and v as the a priori state reflecting the information about the system prior to the measurement. The partial information gained from the measurement of $\mathscr{B}$ is represented by a normal state $w_{\mathscr{B}}$ on $\mathscr{B}$.

According to a theorem by Takesaki,[5] there exists uniquely a positive $T \in \mathscr{B}$, $0 < T \leqslant \lambda^{1/2} I$, such that

$$w_{\mathscr{B}}(B) = v_{\mathscr{B}}(TBT); \qquad \forall B \in \mathscr{B}$$

provided that $w_{\mathscr{B}}$ is majorized by $\lambda v_{\mathscr{B}}$ with $1 \leqslant \lambda$. Using this operator T, we define on $\mathfrak{A}$ the state

$$w(A) = v(TAT); \qquad \forall A \in \mathfrak{A} \tag{1}$$

and call it the *statistical inference from* $w_{\mathscr{B}}$ *relative to the subalgebra* $\mathscr{B}$ *and the a priori state* v.

The interpretation of w as a statistical inference is based on the following properties:

(a) w is an extension of $w_{\mathscr{B}}$.

(b) If $w_{\mathscr{B}} = v_{\mathscr{B}}$, then $w = v$.

(c) If $\mathscr{B} = \{\lambda I\}$, then $w = v$.

(d) If $\mathscr{B} = \mathscr{A} = B(H)$, then $w = w_{\mathscr{B}}$.

Property (a) tells us that the inferred state is compatible with the partial information; (b) that no information can be gained from a measurement that agrees with the a priori information; (c) that no information is obtained if the measurement is trivial; and (d) that a complete measurement determines the state on $\mathscr{A}$ completely. Furthermore, it can be shown[1] that the definition (1) is a generalization of the classical definition of statistical inference, and that the state w maximalizes the entropy within the equivalence class of states that are compatible with $w_{\mathscr{B}}$.[6]

In view of our application, we now consider the special case where $\mathscr{A} = B(H)$ is the von Neumann algebra of all bounded operators in H and $\mathscr{B} = \{E_i\}''$ is the Abelian subalgebra generated by the partition of the identity $\{E_i\}$. As shown in Ref. 1, the inference w of $w_{\mathscr{B}}$ relative to $\mathscr{B}$ and v then assumes the explicit form

$$w(A) = \mathrm{Tr}(WA); \qquad W = TVT; \qquad T = \sum_i \left(\frac{w_{\mathscr{B}}(E_i)}{v_{\mathscr{B}}(E_i)}\right)^{1/2} E_i \tag{2}$$

where W is the density operator of w and V the density operator of v defined by $\mathrm{Tr}(VA) = v(A)$, $\forall A \in \mathscr{A}$.

3. THE MEASURING PROCESS

Let us apply the concept of inference described in Section 2 to the analysis of the quantum mechanical measuring process.

We denote by H_S and H_A the Hilbert spaces of the measured system S and the apparatus A, and by the tensor product $H_S \otimes H_A$ the Hilbert space of the combined system $S + A$. The algebra $\mathscr{A} = B(H_S \otimes H_A)$ contains the observables of $S + A$. The measurement consists in reading, say, n positions of the measuring instrument. Since the apparatus is classical,[2] we assume that these positions correspond to n orthogonal vector states $\psi_i \in H_A$. The projectors Q_i onto ψ_i form a partition of the identity in H_A, and the measured subalgebra in $\mathscr{B} = \{I \otimes Q_i\}''$. If w_i are the probabilities of the instrument readings $i = 1, 2,..., n$ as obtained from a sequence of identical experiments, then the state $w_{\mathscr{B}}$ on $\mathscr{B}$ is uniquely defined by $w_{\mathscr{B}}(I \otimes Q_i) = w_i$.

The a priori state v on $\mathfrak{A}$ contains the information about the system $S + A$ before the measurement, which consists of our knowledge of (a) the observables of S that A is supposed to measure, and (b) the accuracy with which A measures these observables. Mathematically, this amounts to the specification (a) of n orthogonal states $\varphi_i \in H_S$ and the corresponding projectors P_i in H_S, and (b) of a correlation matrix $\epsilon_{ij} = E(Q_j \mid P_i)$, where $E(Q_j \mid P_i)$ is the conditional probability that if S is in the state φ_i, then A is in the state ψ_j. Since the projections $P_i \otimes I$ and $I \otimes Q_j$ commute, we have

$$\epsilon_{ij} = E(Q_j \mid P_i) = v(P_i \otimes Q_j)/v(P_i \otimes I)$$

Next we require that the a priori probability $v(P_i \otimes I) = \alpha$ is independent of i. (If this were not so, the a priori state v would yield some information about the state of S before any measurement is performed!) It then follows from the normalization of v that $\alpha = 1/n$. Hence

$$v(P_i \otimes Q_j) = (1/n)\, \epsilon_{ij} \tag{3}$$

The constraints (3) do not define the state v uniquely. We therefore invoke the "principle of sufficient reason" by defining v as the maximum entropy state under the constraints (3).

Lemma. Within the equivalence class of states satisfying (3), the state defined by

$$v(A) = \mathrm{Tr}(VA); \qquad V = \frac{1}{n} \sum_{k,l} \epsilon_{kl}(P_k \otimes Q_l) \tag{4}$$

has maximum entropy.

Proof. The state v defined by (4) satisfies the constraints (3):

$$v(P_i \otimes Q_j) = \mathrm{Tr}[V(P_i \otimes Q_j)] = \frac{1}{n} \sum_{k,l} \epsilon_{kl} \,\mathrm{Tr}(P_k P_i \otimes Q_l Q_j)$$
$$= \frac{1}{n} \epsilon_{ij} \,\mathrm{Tr}(P_i \otimes Q_j) = \frac{1}{n} \epsilon_{ij}$$

That v is the maximum entropy state can be seen as follows: In the basis $\varphi_i \otimes \psi_j$ of $H_S \otimes H_A$ the constraints read $V_{ij,ij} = (1/n)\, \epsilon_{ij}$. Within the equivalence class of density matrices with these diagonal elements the von Neumann entropy $-\mathrm{Tr}(V \log V)$ has a maximum for the matrix V whose off-diagonal elements vanish.[7] It is therefore sufficient to verify that (4) is diagonal in the basis $\varphi_i \otimes \psi_j$. But this follows immediately from the fact that (4) is a function of $P_k \otimes Q_l$.

With $\mathfrak{A}$, $\mathfrak{B}$, v, and $w_{\mathfrak{B}}$ specified, we now apply the results of Section 2 in order to compute the inferred state w on the algebra $\mathfrak{A}$ of the combined system. Using (2), we obtain

$$w(A) = \mathrm{Tr}(WA); \qquad W = TVT; \qquad T = \sum_i \left(\frac{w_i}{v_i}\right)^{1/2} (I \otimes Q_i)$$

where $w_i = w_{\mathfrak{B}}(I \otimes Q_i)$ and

$$v_i = v_{\mathfrak{B}}(I \otimes Q_i) = \mathrm{Tr}[V(I \otimes Q_i)] = \frac{1}{n} \sum_{k,l} \epsilon_{kl} \,\mathrm{Tr}[(P_k \otimes Q_l)(I \otimes Q_i)]$$
$$= \frac{1}{n} \sum_{k,l} \epsilon_{kl} \,\mathrm{Tr}(P_k \otimes Q_l Q_i) = \frac{1}{n} \sum_k \epsilon_{ki} \,\mathrm{Tr}(P_k \otimes Q_i) = \frac{1}{n} \sum_k \epsilon_{ki}$$

Putting these results together, we find that the density operator W of the combined system becomes

$$W = \sum_{k,l} \frac{\epsilon_{kl} w_l}{\sum_m \epsilon_{ml}} (P_k \otimes Q_l) \tag{5}$$

The state w_S of the measured system S can be derived from the state w of the system $S + A$ as follows: Let $\mathfrak{A}_S = B(H_S)$ be the algebra of observables in H_S. Then the restriction of w to $\mathfrak{A}_S$ is defined by

$$w_S(A) = w(A \otimes I); \qquad \forall A \in \mathfrak{A}_S$$

In terms of density operators, this can be written

$$\mathrm{Tr}_S(W_S A) = \mathrm{Tr}[W(A \otimes I)] \tag{6}$$

where Tr_S is the trace over H_S, and a simple argument shows that (6) has the unique solution $W_S = \mathrm{Tr}_A W$, with Tr_A the trace over H_A. Hence the density operator of the measured system is

$$W_S = \sum_{k,l} \frac{\epsilon_{kl} w_l}{\sum_m \epsilon_{ml}} P_k \tag{7}$$

The inferred state of the apparatus can be derived in a similar fashion by setting $W_A = \mathrm{Tr}_S W$. The result is, as expected,

$$W_A = \sum_l w_l Q_l$$

We conclude this section with a direct verification of (7) for the special case where the system S has been prepared in the maximum entropy state $W_0 = (1/n)\, I$. Since the prepared state contains all the information about the system, the entropy of the inferred state W_S should not be less than the entropy of W_0. We therefore expect $W_S = W_0$.

Observe now that the probability w_l for the instrument reading l in the prepared state W_0 is

$$w_l = \sum_m E(Q_l \mid P_m)\, E(P_m) = \frac{1}{n} \sum_m \epsilon_{ml} \tag{8}$$

where $E(P_m) = \mathrm{Tr}(W_0 P_m) = 1/n$ is the a priori expectation of the event P_m in state W_0. If (8) is inserted into (7), we obtain, as expected,

$$W_S = \sum_{k,l} \frac{\epsilon_{kl}(1/n) \sum_m \epsilon_{ml}}{\sum_m \epsilon_{ml}} P_k = \frac{1}{n} \sum_{k,l} \epsilon_{kl} P_k = \frac{1}{n} \sum_k P_k = \frac{1}{n} I = W_0$$

4. REDUCTION OF STATES

In Section 3 we have derived the inferred state W_S of the measured system S for arbitrary correlations ϵ_{ij}. Let us now consider the "perfect measurement." We show that in this case inference corresponds mathematically to reduction.

In a perfect measurement the states φ_i of S and ψ_j of A are totally correlated. The conditional expectations are therefore

$$\epsilon_{ij} = E(Q_j \mid P_i) = \delta_{ij} \tag{9}$$

If (9) is inserted into (7), we obtain

$$W_S = \sum_k w_k P_k \tag{10}$$

Suppose now that the system S has initially been prepared in the pure state,

$$W_0 = P_\varphi ; \qquad \varphi = \sum_i \alpha_i \varphi_i$$

where P_φ is the projection onto φ. Since the events P_i and Q_j are totally correlated, we have $w_k = \mathrm{Tr}(W_0 P_k) = |\alpha_k|^2$ and (10) reads

$$W_S = \sum_k |\alpha_k|^2 P_k \tag{11}$$

The state (11) has been called the "reduction" of W_0. Reduction has sometimes been attributed to a "collapse" of the prepared state under an acausal physical interference of the observer with the system. As our analysis shows, the collapse of the state can also be understood as a consequence of incomplete information about the system.

REFERENCES

1. J.-P. Marchand, *Found. Phys.* **7**, 35 (1977).
2. J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, 1968), Chapter 11.9.
3. H. Spohn, *Int. J. Theor. Phys.* **15**, 365 (1976).
4. J. M. Jauch, E. P. Wigner, and M. M. Yanase, *Nuovo Cim.* **48B**, 144 (1967).
5. M. Takesaki, *Tomita's Theory of Modular Hilbert Algebras* (Springer Lecture Notes in Math., No. 128, 1970), p. 104.
6. J.-P. Marchand and W. Wyss, *J. Stat. Phys.* **16**, 391 (1977).
7. J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955), p. 379.