Zeitschrift: Helvetica Physica Acta
Band: 37 (1964)
Heft: IV-V

Artikel: The problem of measurement in quantum mechanics
Autor: Jauch, J.M.
DOI: https://doi.org/10.5169/seals-113486

The Problem of Measurement in Quantum Mechanics

by **J. M. Jauch**
University of Geneva and CERN-Geneva

(30. I. 64)

Abstract: This is a new analysis of the measuring process for non-relativistic quantum mechanical systems, in order to clarify the well-known difficulties in the interpretation of this process. The rules of quantum mechanics prescribe two apparently different and unrelated changes of the state of a system under the measuring process. It is shown that the two ways of the change of state vectors can be understood without introducing von Neumann's 'ultimate observer' and without abandoning the linear law of the time evolution of states. Consciousness or even the macroscopic nature of the measuring device is not an essential requirement for a measurement. What is required is only a 'classical' property already present in certain micro-systems. For such a classical system the states fall into classes of equivalent states which cannot be distinguished by any observation on the system. It is shown that the two states obtained in the measuring process are in the same equivalence class. Thus the problem concerning this strange duality which has haunted quantum mechanics from the beginning dissolves into a pseudoproblem.

I. Introduction

During recent years there has been a surge of interest in the fundamental problems of quantum mechanics. Many of these problems were essentially solved during the heroic period of the late twenties when quantum mechanics was discovered. The great success of this theory in its various applications to atomic and nuclear physics has tended to emphasize its pragmatic aspects, and many young physicists have not always appreciated the enormous effort which was expended during the early days in order to obtain a satisfactory *interpretation* of the formalism.

This interpretation has been largely the work of Bohr and his school and one often refers to it as the Copenhagen interpretation. It is based primarily on a careful analysis of the measuring process.

Already in an early paper, Heisenberg[1]) has shown that classically familiar quantities like *position* and *momentum* have only a meaning on the atomic level, insofar as a procedure is known which permits at least in principle the determination of the values of such quantities. In contradistinction to classical mechanics, the measuring process enters in an essential way in the interpretation of the mathematical formalism of quantum mechanics. The uncontrollable influence of the measuring device on the measured object leads to limitations of the accuracy of measurements known as the *uncertainty relations*.

The subsequent elaboration of this aspect of quantum mechanics by Bohr led to the general notion of *complementarity*, which expresses the fundamental limitation imposed on all quantum measurements, obtained by mutually exclusive experimental arrangements with equipment which operates on the classical level of perception[2]).

The result of this analysis shows that BORN's probabilistic interpretation[3]) of the Schrödinger wave function can be carried through consistently. The subsequent success of quantum mechanics leaves little doubt that at least the more elementary (non-relativistic) part of the theory gives a correct description of the physical phenomena.

Yet, there remain certain features in the theory which have paradoxical character and which have been the subject of much controversy[4]). It is a group of phenomena which are often loosely referred to as 'the reduction of the wave packets'. This expresses the fact that there seem to be two entirely different changes for the state-vector of a physical system, one abrupt and not entirely predictable which takes place during the measuring process and the other continuous and causal during all other times.

In the language of the statistical operator these two changes of the state may be expressed as follows[5]): let W be the statistical operator, representing the state of the system, R an observable (self-adjoint operator) with discrete (non-degenerate) spectrum, φ_n a complete system of eigenvectors of R, P_n the projection operator, the range of which is the ray containing φ_n, and H the total Hamiltonian for the system. Then the two kinds of changes are expressed by the formulae

$$W \rightarrow W' = \sum_n (\varphi_n, W \varphi_n) P_n , \tag{1}$$

$$W \rightarrow W_t = e^{-iHt} W e^{iHt} . \tag{2}$$

It is hardly necessary to emphasize the gulf which separates these two processes. Suffice it to recall here that the second change always produces a statistical operator W_t which is unitarily equivalent to W, while the first change in general does not do this.

The central question of the problem of measurement is this: how can these two changes of the state be reconciled? This question is the subject of the present paper.

An obvious attempt at an explanation is the following: the states behave differently because of different situations. In situation (1) the system (S) is in contact with a measuring instrument (M) and only the combined system $S_1 = (S + M)$ behaves according to a Schrödinger equation (2). The difficulty with this explanation is that, in order to determine the state of the system $S + M$, another observer (M_1) is needed who registers this state, and during the process when M_1 observes $S + M$, it is the change (1) which determines the outcome. Thus one has not succeeded in reducing process (1) entirely to process (2), one has only replaced the 'cut' between the system and the observer to the position between $S + M$ and M_1.

One can, of course, continue this process and instead of considering the system $S_1 = S + M$ one may consider the system $S_2 = S + M + M_1$. The state of this new system will evolve according to process (2) but again, in order to determine the state, a third observer M_2 is needed and during the process of observation it is again the process (1) which determines the result.

Thus again one has not eliminated process (1), one has only shifted the 'cut' between system and observer to still another place, and it is obvious that one never will, with this method, succeed in eliminating process (1).

The best that one can hope to do, and this was done by VON NEUMANN[5]), is to show that the result of the measurement on S is independent of the location of the 'cut'. This result enables one to restore the objective character of the measuring process.

The unsatisfactory feature of this explanation is that one needs an 'ultimate observer' M_0 which will appear at the end of an infinite regression, and that one has not succeeded in eliminating process (1) from the theory. This ultimate observer is supposed to be endowed with a nonphysical property, called 'consciousness' and one tries to make the entering of the perception into consciousness responsible for the occurrence of process (1)[6]). We cite from this last reference the following passage which expresses the point of view indicated in the text:

«Ce n'est donc pas une interaction mystérieuse entre l'appareil de mesure et l'objet qui produit dans la mesure l'apparition d'un nouveau ψ du système. C'est seulement la conscience d'un Moi qui se sépare de la fonction $\psi(x, y, z)$ ancienne et constitue une nouvelle objectivité en vertu de son observation consciente en attribuant désormais a l'objet une nouvelle fonction d'onde $u_k(x)$.»

We do not share this point of view. It would seem that it would amount to a tacit admission that a completely objective physical theory is impossible. We would have to choose between the occurrence of two irreconcilable process laws as expressed in (1) and (2), and admit the existence of an entity (consciousness) which stands outside the physical laws. The inadequacy of this 'explanation' has been felt by many physicists and there have been many attempts in the past to replace this peculiar situation by something better[7]). They are usually directed towards finding in the macroscopic nature of the measuring device M the reasons for the occurrence of mixtures at the end of the measuring process. However, in a recent paper on the measuring process, WIGNER[8]) has shown that such attempts are inconsistent with the linear laws of quantum mechanics as expressed for instance in Equation (2). But he, too, remains uneasy about this 'strange dualism' as expressed in the two Equations (1) and (2) and he even considers the possibility of abandoning the linear equation of motion for the state vector[9]).

The unsatisfactory feature of VON NEUMANN's infinite regression and the 'ultimate observer' is made evident by the simple device of including the 'ultimate observer' with consciousness and all in the measuring apparatus M. In order that the perception enters a consciousness it is then no longer necessary that an 'ultimate observer' outside the system makes the observation which furnishes the reduction of the wave packet. The contemplation of this example could easily lead to the conclusion that the laws of quantum mechanics are not applicable for measuring devices with consciousness[10]).

We shall in this paper propose an interpretation of the measuring process which disposes both of VON NEUMANN's 'ultimate observer' and the need for abandoning the linear law of the evolution of states. This interpretation is based on a more careful analysis of the notion of the 'state' of a quantum-mechanical system.

It was emphasized especially by BOHR that measurements are made with equipment which operate on the classical level of perception. This aspect is lost in von Neumann's analysis of the measuring process, where system and observer are treated alike. Process (1) is only needed if the instrument of measurement M is a classical apparatus. In that case the states cannot be distinguished as sharply as they can be represented mathematically. This fact expresses itself by the appearance of an *equivalence relation* between quantum mechanical states. Two states are equivalent if the classical observer cannot distinguish them. It makes then sense to refer to classes of equivalent states as 'macrostates'.

A macrostate can then be represented by any one member of the class. It will be shown that the resulting freedom of choice is just what is needed to reconcile the two different processes (1) and (2). Thus our interpretation of the measuring process restores unity of the process law by heeding Bohr's insistence of the ultimately classical aspect of all measurements.

Our analysis has other implications which we shall not pursue here in detail. It throws a new light on a number of so-called paradoxes such as Schrödinger's cat[11]) and the paradox of Einstein, Podolski and Rosen[12])[13]).

In order to give a unified and self-contained presentation, we shall repeat many well-known things which are not controversial, interspersed with some comments which we believe are new.

In Section II we introduce the notion of the state of a quantum mechanical system. Section III is devoted to the question of the interpretation of the state. In Section IV we introduce the theory of the equivalence classes of states. The following Section V contains the well-known formalism of the union and separation of systems. The only novelty here is a co-ordinate-free definition of the tensor product. The loss of 'Anschaulichkeit' in this procedure is amply compensated by the resultant elegance and simplicity in the discussion of interacting systems. Section VI finally gives the new interpretation of the measuring process and the resultant reconciliation of the two processes (1) and (2).

II. The State of a Quantum-Mechanical System

We shall assume that we are dealing with a quantum mechanical system (that is, a system for which the departure from classical physics is an essential feature) that we can in principle sufficiently isolate this system from the rest of the world and that we can make controlled experiments with it. Such systems are for instance a spin in a paramagnetic substance, an isolated nucleus, a system of electrons, a photon, or a neutrino. We shall further assume that at every instant the system is in a definite *state*. What does this mean?

For the following it is rather important to understand the answer to this question, so we shall discuss it rather carefully, perhaps more so than it is usually done and necessary for most purposes.

There are three aspects of the state of a system which concern us here, its preparation, its determination and its mathematical description. We shall take these three aspects in turn.

Every state of a physical system is the result of a preparation of the system. A preparation is a series of manipulations with physical equipment which affect the system under consideration. The difficulty with this definition is that it is not immediately obvious whether certain physical conditions will affect a state or not. Thus for instance the presence of a static electric field does (in a very good approximation) essentially nothing to a magnetic ion. However if the field is oscillating the state is markedly affected. The state of the photons emitted from a discharge tube could depend on the intensity of the discharge, but they do not. We are thus led to the notion of *relevant* conditions for the preparation of a state. It is an empirical fact that certain conditions will play no role in the preparation of a state, while others do, the ones that do are the relevant ones. We shall have to accept this fact without further analysis and without being able to say, in each given case, what these conditions are.

We summarize: a state is the result of a series of physical manipulations on the system which constitute the preparation of the state. Two states are identical if the *relevant* conditions in the preparation of the state are identical.

Almost the entire difficulty of this notion is hidden in the word 'relevant'. What is relevant and what is irrelevant is part of the physical law and this is empirically given and so is not known *a priori*.

We shall not dwell on this aspect any further except to point out that the *epoch* at which the state is prepared is usually irrelevant. This is very fortunate for the possibility of determining the state. A state can be determined by measuring every observable quantity of the system. In general the determination of an observable is itself a relevant condition and will therefore modify the state. Thus in order to determine a state it is necessary to be able to repeat the preparation of the state under identical relevant conditions. The outcome of such measurements is not always certain but has instead only a definite probability distribution which can be determined to any degree of accuracy by repeating the measurements a sufficient number of times.

We summarize: the *determination* of a state always requires the preparation of an ensemble of identical systems under identical relevant conditions. The state is determined if the probability distribution of every observable quantity is measured on this ensemble[14]).

Turning now to the mathematical description of the state we see that a state is completely determined if we give a probability function on every observable. Since an observable can in general assume many different values, it is convenient to introduce a special class of observables which can assume only two values. It is not hard to see that the measurement of every observable can be reduced to the measurement of a sufficient number of such special observables. We call such observables yes-no experiments or propositions and we fix the two values arbitrarily to be represented by 1 or 0 (true or false).

In the usual form of quantum mechanics such observables are described by projection operators and the above statement about general observables has its mathematical counterpart in the spectral theorem of self-adjoint operators.

The probability function which describes a physical state is a more general object than the usual (classical) probability function. This latter is always defined on a Boolean lattice of classes of subsets and it is one of the basic facts of quantum mechanics that the set of all the questions is not a Boolean lattice, in fact it is not even a modular lattice[15]). We stress here that this statement is empirically given and involves essentially no assumption since the lattice operations have a direct physical interpretation[16]). The probability function retains on every Boolean sublattice of propositions the properties of an ordinary probability.

A state is thus mathematically given if we are given a real valued function $p(E)$ on the projection operators representing propositions, with the following properties

$$0 \leq p(E) \leq 1 , \quad p(0) = 0 , \quad p(I) = 1 . \tag{a}$$

$$\text{If} \quad E F = 0 \quad \text{then} \quad p(E + F) = w(E) + w(F) . \tag{b}$$

The first of these is obvious, the second expresses the additivity of the probability function on mutually exclusive questions. Two questions E and F are mutually exclusive if whenever E is true then F is false, and vice versa. The additivity property can then be easily inferred for instance from the frequency interpretation of the probability function.

A remarkable mathematical thematical theorem[17]) asserts that a functional $p(E)$ defined on all the projections E of a (separable) Hilbert space which satisfies conditions (*a*) and (*b*) must be of the form*)

$$P(E) = Tr(W E) , \tag{3}$$

where W is a linear operator which satisfies the conditions

$$\left.\begin{array}{ll} \text{i)} & W^* = W , \\ \text{ii)} & W^2 \leq W , \\ \text{iii)} & Tr W = 1 \end{array}\right\} \tag{4}$$

The theorem of GLEASON quoted here is useful to show that the representation of quantum mechanical states is much less arbitrary than it is customarily assumed (or at least presented). It shows that if one tries to generalize states, one would have to do it in the sense of footnote[14]) or one would have to assume that the observable projections do not generate an algebra of type I[18]). For the elementary systems of non-relativistic quantum mechanics the applicability of GLEASON's theorem is certain.

The operator W is VON NEUMANN's density operator. Every W which satisfies condition (4) has a discrete spectrum. Its spectral resolution has thus the form

$$W = \sum_n \lambda_n P_n , \tag{5}$$

where P_n is a projection operator with one-dimensional range. (If some eigenvalues are degenerate we repeat them.) The eigenvalues λ_n satisfy then

$$\left.\begin{array}{ll} \text{i)}' & \lambda_n^* = \lambda_n , \\ \text{ii)}' & 0 \leq \lambda_n \leq 1 , \\ \text{iii)}' & \sum_n \lambda_n = 1 . \end{array}\right\} \tag{4'}$$

*) Actually the additivity property (b) is required to hold even for an infinite sequence of projections.

If $W^2 = W$ then W is a projection operator of one-dimensional range. In that case we call the state *pure*. In all other cases it is called a *mixture*. If W is a pure state then any unit vector φ in the range of W may be taken as an alternate representation of W. Such a φ is the Schrödinger wave function corresponding to this state. Clearly φ is only determined up to a numerical factor of magnitude 1. For pure states (3) becomes

$$p(E) = (\varphi, E\,\varphi)\,. \tag{3'}$$

If φ_n are the eigenstates of the projections P_n we shall later on occasionally refer to the state given by (5) as the 'mixture of the states φ_n'.

Most of these things are well-known and we have mentioned them here merely in order to have a well-defined frame of reference within which we shall analyze the delicate problem of measurement in quantum mechanics.

Here we want to add a few remarks which serve to emphasize some points which will be useful in the following. First we mention that the definition of states which we have given here is also applicable to classical systems, although it is a little more general than usual. The difference with respect to quantum systems is only this: every state is a mixture with the exception of those with a probability measure concentrated on one single point in phase space, and every proposition is compatible with every other one. A general state in classical mechanics will therefore be described by a general probability measure in phase space. The relevant conditions in the preparations of the state involve certain *a priori* probabilities in the distribution of physical conditions.

Thus, for instance, if throwing a die is considered the preparation of a state for the die, the discrete six different configurations are for a normal die under normal conditions distributed with probability 1/6. The identical relevant conditions are the physical situations with the same *a priori* probabilities for the initial conditions.

After these remarks we discuss now an example which illustrates some of the relevant conditions which may occur in the preparation of a quantum mechanical system. We consider two particles of spin 1/2 constrained to move on a straight line. They interact only upon contact and in such a manner that each particle is completely reflected so that after collision with the other it reverses its motion. The interaction is however assumed to be spin-dependent so that the spin states can change during the impact. It is only the spin degrees of freedom which interest us. Each particle has exactly two states

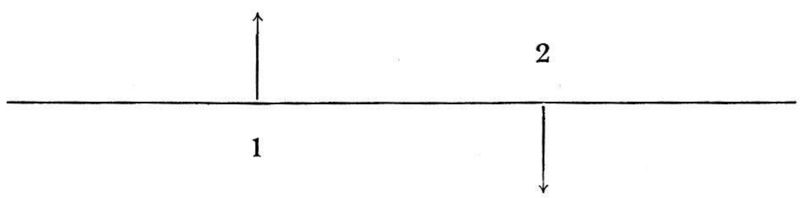

Fig. 1
Collision of two spins with initial state $u_1\,v_2$

which we designate with u_1 and v_1 for particle 1 und with u_2 and v_2 for particle 2.

If the interaction conserves total spin we find that triplet and singlet states do not mix and furthermore the states $u_1\,u_2$ and $v_1\,v_2$ do not change. Thus if we denote by

$$\psi_1 = \frac{1}{\sqrt{2}}\,(u_1\,v_2 + v_1\,u_2)\,,$$

one of the triplet states, and by

$$\psi_0 = \frac{1}{\sqrt{2}} (u_1 v_2 - v_1 u_2)$$

the singlet state and with ψ_1' and ψ_0' the corresponding quantities after collision we have (after suitable adjustment of an over-all phase factor)

$$\psi_1' = e^{i\delta} \psi_1 ,$$
$$\psi_0' = e^{-i\delta} \psi_0 .$$

It follows from this that, if the initial state is (as indicated in Fig. 1) $\Psi = u_1 v_2$, then the final state after scattering is

$$\Psi' = \cos\delta \, u_1 v_2 + i \sin\delta \, v_1 u_2 .$$

what we have described so far is the ordinary theory of triplet and singlet scattering. But now we look at this experiment in a different way. We consider it as the preparation of a state for particle 1. The relevant conditions which we impose on particle 1 are the following:

a) prepare particle 1 in state u_1;
b) prepare particle 2 in state v_2;
c) bring them to collision.

We consider thus an ensemble of particles of type 1, all members of which were prepared under the conditions a), b) and c). Every member of such an ensemble is in a well-defined state which we can calculate easily and measure in principle. An elementary calculation, which we need not reproduce here, shows that the state in question is a mixture, described by the density matrix

$$W = \begin{pmatrix} \cos^2\delta & 0 \\ 0 & \sin^2\delta \end{pmatrix} .$$

We have here thus an explicit example where a pure state (u_1 is changed into a mixture by subjecting the system to a complete set of certain relevant conditions.

Let us now prepare a different state. The new state shall be prepared by imposing as before the conditions a), b) and c), and a fourth one which might be expressed as:

d) particle 1 must be in coincidence with spin up of particle 2.

This means in the beam of particles 1 we select only those which are associated with particles in beam 2 with spin up. We can prepare this state by constructing an automatic shutter for particle 1 which is activated by a spin measuring equipment for particle 2 and traps all those particles of kind 1 which are associated with spin of particle 2 down.

The conditions a), b), c), and d) define a new ensemble of particles 1 which is clearly a sub-ensemble of the previous one. In this new ensemble we can again determine the state by measurement and calculation. There is no reason why this state should be the same as in the previous case, since a relevant condition d) has been added in its preparation. And indeed it is not. The new state is pure and is given by v_1, as one might almost guess without calculation.

We shall encounter exactly the same situation in the measuring process. The mere observation of a property may not interfere with a state if it is a mixture but it might if it is used for the preparation of a new state with different relevant conditions.

III. Interpretation of States

Until now we have described the preparation, determination and mathematical description of states. We shall face now the more difficult question of the *interpretation* of states. We can do this best if we compare the classical with the quantum system.

We have already remarked that one definition of state is also applicable to classical systems and we have pointed out that even in the classical case a general state will be a probability measure in phase space. There is therefore nothing unusual about the occurrence of probabilities in the description of states.

Yet there is a profound difference in the interpretation of these probability functions in the case of quantum mechanics and classical mechanics. This difference is often somewhat loosely described by the statement that in quantum mechanics the determination of an observable by measurements inevitably involves an uncontrollable interaction of the measured object with the measuring device and that this interaction produces a distribution of values of the measured quantities in accordance with the numerical expectation values calculated for this quantity in the particular state. It was precisely this analysis which led to the physical interpretation of the uncertainty relation[1]).

It was, however, pointed out repeatedly by EINSTEIN[12]) that there are quantum mechanical states for which this restriction does not apply. In the quoted paper it was shown that there exist states which predict only a probability outcome for certain variables. Yet measurements of such variables can be made without in any way interfering with the state in question. In Reference [12]) this was accomplished by separating two interacting systems and by carrying out measurements on one, in order to determine certain quantities in the other. A closer analysis of this situation reveals that the states which have such properties are precisely the mixtures with respect to the spectral projections of the measured quantity. Let us illustrate this in a particular situation of sufficient generality to show that the feature is a general one.

Let R be the measured quantity with a spectrum which we assume to be discrete and non-degenerate. Let φ_n be the eigenvector of R and P_n the projection with range φ_n. We consider a general state given by a density operator W. We ask, under what condition on W does the process (1) leave W unchanged. The condition on W is thus

$$W' = W = \sum_n (\varphi_n, W \varphi_n) P_n .$$

It follows from this equation that W is of the form

$$W = \sum_n p_n P_n , \tag{6}$$

with $0 \leq p_n \leq 1$. The p_n are otherwise arbitrary. Since for any such W we have also

$$(\varphi_n, W \varphi_n) = (\varphi_n, P_n \varphi_n) = p_n ,$$

we see that this condition is also sufficient. Thus the spectral projections of a statistical mixture can, if they are measurable, be determined without in any way interfering with the state of the system.

It is useful to introduce a special terminology to distinguish such properties from the others which cannot be so determined. Thus Einstein attributes to such properties in a given state an 'element of reality'. It makes sense to attribute to each individual system in such a state *one* of the alternatives of the statistical mixture even though the state which resulted from some preparation allows only the prediction of probabilities. We could, if we wanted to, determine this property without affecting the state.

This is, of course, exactly the classical situation. In classical systems every state allows such a realistic interpretation of probabilistically described events.

We shall use a shorter terminology which we find adequate to distinguish the two types of properties in a given state. We shall use the word 'event' to denote a physical property which can be observed without interfering with the state of a system. When such an event has been observed we say: the event is a 'datum'. Thus in a state of the form (6), every individual system from an ensemble of identically prepared systems realizes one of the events P_n, and this independently whether a subsequent observation has raised the event to the level of a datum or not.

Suppose now we consider a classical system in a given state, for instance a die thrown under certain specified conditions. In general the specification is so that the outcome is a state which allows only probability statements for the observation of subsequent events. In the example of the die under usual specifications, each side will occur with probability 1/6. However, we know that classical systems are such that they permit the addition of further relevant initial conditions until the state is free from dispersion in its predicted outcome of future measurements. For instance, we could specify the initial conditions for the die with such precision that the outcome of the throw is determined by these conditions. In this case the probability function would have the value 1 for that particular event and zero for the others.

If we examine the same question for quantum systems, we find that it is not always possible to add further relevant conditions in order to reduce the dispersion of a state. For pure states, this is a familiar feature. What is perhaps less well-known is that this situation may occur even for mixtures. This is a profound difference between classical and quantum systems alluded to above and this is the origin of most of the paradoxical features of quantum mechanics.

In order to demonstrate this, we refer to the example of the preceding Section, which we have discussed for just this reason. If we consider the conditions a), b), and c) of this example, the conditions preparing the state of the spin 1, we find that these conditions determine only a statistical mixture of this spin. There is no further relevant initial condition known, which one could add to this preparation and which would determine a pure state of the spin 1. We could, of course, always, *after* the preparation of the state, add a further condition which selects from the statistical mixture a pure state (for instance condition d) in the example of Section II). What we cannot do is add such a condition *before* the preparation of the state.

It is here where we encounter the essentially irreducible statistical element of quantum mechanics in its full force. Since we attributed with good reasons to the alternatives of a statistical mixture an element of reality which permits us to regard them as 'events', we find that *different* events may occur under *identical* relevant conditions. We are making here of course an *empirical* statement. We can only state that such conditions are not *known* to exist.

This finding seems to be in contradiction with the principle of sufficient reason which says roughly that identical causes produce identical effects. If one wants to escape this conclusion without being in conflict with empirical facts, the only possibility left seems to be to assume the existence of hidden (non-observable) determining variables, which are correlated with the events of the statistical mixtures we have been discussing. Unfortunately this way out of the above conclusion is barred by the fact, demonstrated first by VON NEUMANN, that such hidden variables are incompatible with the other observable consequences of quantum mechanics[19]).

EINSTEIN, PODOLSKI and ROSEN have summarized this situation with the statement that the description of a state with the wave function is *incomplete* for individual systems. The critique of BOHR[13]) does, strictly speaking, not invalidate this conclusion, it merely shows that the properties which can be considered as events are not in contradiction with the uncertainty relations. Indeed for the determination of such properties which violate the uncertainty relations one would need for all cases mutually exclusive physical situations which would necessarily modify the state.

We shall take here the point of view that the ultimate arbitration in this vexing question will not come from philosophical principles but from experience. Whether there are relevant conditions which will correlate with undetermined alternatives in a mixture will and can be decided by experience. All indications are so far that in some cases there are none and that the description of the state is not always capable of specifying all the real physical events of the individual system.

We shall accept this verdict, in the same tentative spirit as one must accept any inductively obtained theory.

We emphasize once more that the difficulty in the interpretation which we are discussing here appears only for mixtures and only if the projections of the mixture are measurable quantities. It is therefore of some importance to know under what conditions we can ascertain when a state is a mixture and when it is pure.

In principle this is always possible if we measure a sufficiently large number of suitable observables. For instance if all projections were observables there would be no difficulty in deciding whether a given state is a mixture or not. However, if not all self-adjoint operators are observables, then the nature of the state can in general not be decided by measurements alone. It becomes then partly ambiguous.

The general discussion of this point will be deferred to the next Section.

IV. Equivalent States

We shall now examine the point mentioned at the end of the last Section concerning tha states in a system with an incomplete set of observables.

Let $\mathfrak{S}$ be the system of observables of a physical system. This is a set of self-adjoint operators. If $A \in \mathfrak{S}$ is an observable then it is operationally meaningful to require that not only A but also A^2 is an observable, since an observation of A which yielded the value λ, is at the same time an observation of the quantity A^2, giving the value λ^2. More generally, if $u(\lambda)$ is a Borel function of the real variable λ, then $u(A)$ is also an observable. If A has the value a then $u(A)$ has the value $u(a)$. From this follows that all the spectral projection operators of A are observables too. Still more generally, if A_i is any set of mutually commuting operators in $\mathfrak{S}$ then any self-adjoint operator X affiliated to the von Neumann algebra generated by A_i is also an observable. The operational justification for this statement is only slightly more complicated than in the example just discussed.

We shall say two states W_1 and W_2 are *equivalent* with respect to the system of observables $\mathfrak{S}$ if

$$Tr\, A\, W_1 = Tr\, A\, W_2$$

for all $A \in \mathfrak{S}$. For two equivalent states we shall write $W_1 \sim W_2$. One verifies without difficulty that this relation is indeed an equivalence relation. Physically, the relation $W_1 \sim W_2$ means that the two states W_1 and W_2 cannot be distinguished by any measurement whatsoever with observables from the system $\mathfrak{S}$.

Since equivalent states cannot be distinguished by any measurement of observables from $\mathfrak{S}$ the association of statistical operators with physical states of the system becomes ambiguous. In order to restore the one-to-one correspondance of physical states with the mathematical description of states, one must introduce the class of equivalent states as the appropriate description of the physical states. One can refer to such classes as *macrostates*, while each individual member in the class could be designated as a *microstate*. Zhe operation of the mixture of states can be transferred from the microstates to the macrostates. This means if W_1 and W_2 are two different microstates then $W = \lambda_1 W_1 + \lambda_2 W_2$ with $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_1 + \lambda_2 = 1$ determines a microstate. The class of states equivalent to W is independent of the representatives W_1 and W_2 in the equivalence classes of W_1 and W_2. This statement may be summarized with the following theorem:

If $W_1 \sim W_1'$ and $W_2 \sim W_2'$ and

$$W = \lambda_1 W_1 + \lambda_2 W_2 ,$$

$$W' = \lambda_1 W_1' + \lambda_2 W_2' ,$$

then $W \sim W'$.

For the proof of this theorem we remark that equivalence of two states is established if we can show the equality of expectation values for observable projections only. The expectation value for general observables follows then from the spectral theorem. Thus we need only verify that $W_1 \sim W_1'$ and $W_2 \sim W_2'$ implies $Tr\, E\, W = Tr\, E\, W'$ for all $E \in \mathfrak{S}_p$. Here $\mathfrak{S}_p$ designates the subset of projections in $\mathfrak{S}$. Since for all $E \in \mathfrak{S}_p$

$$Tr\, E\, W = \lambda_1\, Tr\, E\, W_1 + \lambda_2\, Tr\, E\, W_2 = \lambda_1\, Tr\, E\, W_1' + \lambda_2\, Tr\, E\, W_2'$$
$$= Tr\, E(\lambda_1 W_1' + \lambda_2 W_2') = Tr\, E\, W'$$

the theorem is verified.

Due to this theorem it is possible to transfer the operation of mixture to the equivalence classes. Let us denote by $[W]$ the class of microstates which are all equivalent to W. The statement just proved allows us to define a mixture of macrostates by the formula

$$[W] = \lambda_1[W_1] + \lambda_2[W_2] \equiv [\lambda_1 W_1 + \lambda_2 W_2] . \tag{7}$$

A simple corollary of this theorem is the following statement:

If $W_1 \sim W_2$ then any mixture $\lambda_1 W_1 + \lambda_2 W_2$ is in the equivalence class of $[W_1] = [W_2]$.

A macrostate $[W]$ is said to be pure if it cannot be represented as a mixture of two or more other macrostates.

We shall now examine once more the notion of 'event' introduced in the previous Section in order to describe physical properties which can be determined without interfering in any way with the state. The preliminary discussion of the preceding Section can now be sharpened if we carry it through in the context of the notion of equivalent states. Let W be a microstate, $[W]$ the equivalence class to which it belongs, W^E the microstate after the measurement of the observable projection E, and $[W^E]$ its equivalence class.

The first question arises whether the class $[W^E]$ is independent of the representative $W \in [W]$. If this is the case, we can transfer the change of the state under measurements to the classes. This property can be expressed in the formula

$$[W^E] = [W]^E . \tag{8}$$

In order to obtain the necessary and sufficient condition which guarantees this property, we recall that the measurement of E changes the state W into

$$W^E = E\, W\, E + E'\, W\, E' ,$$

where we have introduced the notation $E' = I - E$. Let now $W_1 \sim W_2$. We ask under what conditions does it follow that $W_1^E \sim W_2^E$? In order that this be true we must have

$$Tr\, F(E\, W_1\, E + E'\, W_1\, E') = Tr\, F(E\, W_2\, E + E'\, W_2\, E')$$

for all $F \in \mathfrak{S}_p$. By using the invariance of the trace under cyclic permutations, we change this into

$$Tr(E\, F\, E + E'\, F\, E')\, W_1 = Tr(E\, F\, E + E'\, F\, E')\, W_2$$

for all $F \in \mathfrak{S}_p$. This is true for all equivalent pairs W_1 and W_2 if and only if

$$F^E \equiv E\, F\, E + E'\, F\, E' \in \mathfrak{S} . \tag{9}$$

Thus we find that the macrostates are left intact under the measuring process if and only if F^E is an observable for all pairs of observable projections E and F. Only if this condition is satisfied, is the meaning of macrostate physically useful. We note here that for classical systems $\mathfrak{S}$ is Abelian and hence (9) is always satisfied. We must ask whether there are macrostates which are left invariant under all measurements with observables in $\mathfrak{S}$. The condition for this is that $[W]^E = [W]$ or, $W^E \sim W$ for all $W \in [W]$. This means, for all $F \in \mathfrak{S}_p$ and all $E \in \mathfrak{S}_p$, we must have

$$Tr\, F\, W^E = Tr\, F\, W = Tr\, F^E\, W\,.$$

A sufficient condition for this to hold is

$$F^E \equiv E\,F\,E + E'\,F\,E' = F \tag{10}$$

for all E, $F \in \mathfrak{S}_p$. If it is satisfied then every state is left invariant under measurements. Since the invariance under measurements is a property of (idealized) classical systems, we might call a state which is left invariant under all observations a *classical state*. We note that condition (10) is satisfied for classical systems. We have thus established that for classical systems every state is classical[20]).

In the discussion of the measuring process in the rest of this paper we shall not need the notion of equivalence in its most general setting. We shall in fact only need it for classical systems. For this case we summarize here the main conclusion: for classical systems the equivalence classes of states are invariant under observation.

V. The Union and Separation of Physical Systems

In this section we shall review the theory of the union and separation of quantum mechanical systems. Although the formalism which describes these processes is well known, we shall present it here once more rather concisely in the co-ordinate free formalism, dwelling particularly on some points which are relevant for the measuring process and making some comments which are new.

If two independent systems I and II are united into a single system I + II we represent the states of the combined system in the product space $\mathfrak{H}^{\mathrm{I}} \otimes \mathfrak{H}^{\mathrm{II}}$ of the state vector spaces $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$ of the component systems. Since this is rather important for the following we shall give a sketch of the reasoning which justifies this assumption.

If the two systems I and II are sufficiently simple (for instance no superselection rules) then we can assume that the system of all observables generates the irreducible algebra of all bounded operators. If the systems I and II are combined into a larger system I + II then (again under the assumption of no superselection rules) the observables pertaining to I are no longer irreducible but they generate only a factor $\mathfrak{F}$. Similarly, the operators of system II generate a factor, the commutant $\mathfrak{F}'$ of $\mathfrak{F}$. Under the starting assumption both $\mathfrak{F}$ and $\mathfrak{F}'$ contain minimal projections. They are therefore of type I. Now according to well-known theorems every such factor defines a tensor product of two Hilbert spaces $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$, such that the factor $\mathfrak{F}$ is isomorphic to the irreducible algebra in $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{F}'$ is isomorphic to the irreducible algebra in $\mathfrak{H}^{\mathrm{II}}$. In this manner one is led to the tensor product for the combined system I + II.

Another more physical reasoning could be based on a discussion of the entropy of states. The negative entropy divided by k of a system in state W is $Tr\, W \ln W$. The entropy being an extensive quantity should behave additively under the combination of two different systems. This is the case if we define the state of the combined system by $W = W^{\mathrm{I}} \otimes W^{\mathrm{II}}$, because one verifies easily that indeed

$$Tr(W^{\mathrm{I}} \otimes W^{\mathrm{II}}) \ln (W^{\mathrm{I}} \otimes W^{\mathrm{II}}) = Tr^{\mathrm{I}}\, W^{\mathrm{I}} \ln W^{\mathrm{I}} + Tr^{\mathrm{II}}\, W^{\mathrm{II}} \ln W^{\mathrm{II}}\,,$$

where Tr^{I} and Tr^{II} denote the trace with respect to the space $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$ respectively.

For the following it is desirable to have a co-ordinate free definition of the tensor product since none of the results depend on the special choice of the reference systems. Furthermore, this formulation is in some ways simpler than the co-ordinate dependent formulation, although it is of course more abstract. In the following we shall expound very briefly the theory of the tensor product emphasizing especially those aspects which will be essential for the discussion of the union and separation of systems. Almost all that we need is contained in Section VI.2 of Reference [5]).

The tensor product is defined in the following way: let $\mathfrak{H}^{\mathrm{I}}$, $\mathfrak{H}^{\mathrm{II}}$ be two Hilbert spaces, the tensor product $\mathfrak{G} = \mathfrak{H}^{\mathrm{I}} \otimes \mathfrak{H}^{\mathrm{II}}$ is a Hilbert space, together with a bilinear mapping φ of the (topological) product $\mathfrak{H}^{\mathrm{I}} \times \mathfrak{H}^{\mathrm{II}}$ into $\mathfrak{G}$, such that

i) the set of all vectors $\varphi(f_1, f_2)$ $(f_1 \in \mathfrak{H}^{\mathrm{I}}, f_2 \in \mathfrak{H}^{\mathrm{II}})$ span $\mathfrak{G}$;

ii) $(\varphi(f_1, f_2), \varphi(g_1, g_2)) = (f_1, g_1)\,(f_2, g_2)$ for all $f_1, g_1 \in \mathfrak{H}^{\mathrm{I}}$, $f_2, g_2 \in \mathfrak{H}^{\mathrm{II}}$.

For complex Hilbert spaces the tensor product always exists and it is unique in the following precise sense:

If $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$ are two Hilbert spaces and $\mathfrak{G}$ and $\mathfrak{G}'$ two different tensor products and φ and φ' the associated bilinear mappings of $\mathfrak{H}^{\mathrm{I}} \times \mathfrak{H}^{\mathrm{II}}$ into $\mathfrak{G}$, $\mathfrak{G}'$ respectively, then there exists a unique isometric operator U with domain $\mathfrak{G}$ and range $\mathfrak{G}'$ and such that

$$U\, \varphi(f_1, f_2) = \varphi'(f_1, f_2)\,.$$

The existence of the tensor product is non-trivial. It is known not to exist for quaternion Hilbert spaces, a major obstacle to a quaternion quantum mechanics[21]). The uniqueness is very important since it shows that the description of the compound system is uniquely determined by that of the component systems (unitary equivalence does not matter!).

An explicit construction of the tensor product for complex spaces can be given as follows. Let $\mathfrak{G}$ be the set of all conjugate linear transformations ϕ from $\mathfrak{H}^{\mathrm{II}}$ into $\mathfrak{H}^{\mathrm{I}}$ which satisfy the following properties

$$\phi(g_2 + h_2) = \phi\, g_2 + \phi\, h_2\,,$$
$$\phi(\lambda\, g_2) = \lambda^* \phi\, g_2\,,$$

for all $g_2, h_2 \in \mathfrak{H}^{\mathrm{II}}$ and all complex λ and

$$\|\phi\|^2 \equiv \sum_r \|\phi\, \psi_r\|^2 < \infty\,.$$

The last expression is to be evaluated for any complete orthonormal system $\{\psi_r\}$ in $\mathfrak{H}^{II}$. It is independent of the choice of that system. The set $\mathfrak{G}$ of these transformations form a Hilbert space. Moreover, there exists a bilinear mapping from $\mathfrak{H}^I \times \mathfrak{H}^{II}$ into $\mathfrak{G}$ by setting $\phi = \varphi(f_1, f_2)$, where

$$\phi\, g_2 = (g_2, f_2)\, f_1 .$$

We denote this particular ϕ with $\phi = f_1 \otimes f_2$. One finds easily that $\| \phi \| = \| f_1 \| \, \| f_2 \|$. To every $\phi \in \mathfrak{G}$ we can associate uniquely another conjugate linear mapping $\phi^\#$ from $\mathfrak{H}^I$ into $\mathfrak{H}^{II}$ by setting

$$(f_1, \phi\, f_2) = (f_2, \phi^\# f_1) .$$

The correspondance $\phi \to \phi^\#$ satisfies the following rules

$$\phi^{\#\#} = \phi \,, \quad \| \phi^\# \| = \| \phi \| \,, \quad (\phi + \psi)^\# = \phi^\# + \psi^\#, \quad (\lambda\, \phi)^\# = \lambda\, \phi^\# .$$

If $\phi = f_1 \otimes f_2$, so that $\phi\, g_2 = (g_2, f_2)\, f_1$, then it follows from the definition of $\phi^\#$ given above that

$$(g_1, \phi\, g_2) = (g_2, \phi^\# g_1) = (g_1, f_1)\,(g_2, f_2) \,,$$

so that

$$\phi^\# g_1 = (g_1, f_1)\, f_2 \equiv f_2 \otimes f_1 .$$

Thus we may write

$$(f_1 \otimes f_2)^\# = f_2 \otimes f_1 .$$

If A_1 is a bounded linear operator in $\mathfrak{H}^I$ and A_2 a similar operator in $\mathfrak{H}^{II}$ we can construct a bounded linear operator $A_1 \otimes A_2$ in $\mathfrak{G}$ by setting for every element $\phi = f_1 \otimes f_2$

$$(A_1 \otimes A_2)\, f_1 \otimes f_2 \equiv A_1 f_1 \otimes A_2 f_2 .$$

From this formula we conclude easily that for every $\phi \in \mathfrak{G}$

$$(A_1 \otimes A_2)\, \phi = A_1\, \phi\, A_2^* .$$

This formalism is especially suited for expressing the effect of the union and separation of systems on the states. This we shall do now. Let W^I be the statistical operator representing a state of system I, and W^{II} the statistical operator for a state of system I. Then the statistical operator of the joint system I + II is given by $W = W^I \otimes W^{II}$.

The converse problem: given W for the joint system, find W^I and W^{II} for the component systems, is more complicated. Since its solution is essential for understanding the measuring process, we shall explain it in some detail. Consider first the case that W represents a pure state (the only case which will be of importance in the application we intend to make); W is then a projection of one-dimensional range and denote by ϕ a unit vector in the range of W. This ϕ is then the state vector representing the state of the joint system.

In order to find the states of the component systems, we must appeal to the physical interpretation. The statistical operators W^{I} and W^{II} are defined by the property that every observable A^{I} referring only to system I and every observable A^{II} referring only to system II must have expectation values

$$Tr^{\mathrm{I}}\,(A^{\mathrm{I}}\,W^{\mathrm{I}}) \quad = Tr\,(A^{\mathrm{I}} \otimes I)\,W\,,$$
$$Tr^{\mathrm{II}}(A^{\mathrm{II}}\,W^{\mathrm{II}}) = Tr(\mathrm{I} \otimes A^{\mathrm{II}})\,W\,.$$

Here Tr^{I} refers to the trace in the space $\mathfrak{H}^{\mathrm{I}}$ and Tr^{II} refers to the trace in the space $\mathfrak{H}^{\mathrm{II}}$. More explicitly, we require that for every complete orthonormal system $\{\varphi_r\}$ in $\mathfrak{H}^{\mathrm{I}}$ and a similar system $\{\psi_s\}$ in $\mathfrak{H}^{\mathrm{II}}$

$$\sum_r (\varphi_r,\, A^{\mathrm{I}}\,W^{\mathrm{I}}\,\varphi_r) = \sum_{r,s} (\varphi_r \otimes \psi_s,\, (A^{\mathrm{I}} \otimes \mathrm{I})\,W\,\varphi_r \otimes \psi_s)$$

and

$$\sum_s (\psi_s,\, A^{\mathrm{II}}\,W^{\mathrm{II}}\,\psi_s) = \sum_{r,s} (\varphi_r \otimes \psi_s,\, (\mathrm{I} \otimes A^{\mathrm{II}})\,W\,\varphi_r \otimes \psi_s)\,.$$

The first formula says that if A^{I} is measured on the system I alone one obtains the same result as if one measures $A^{\mathrm{I}} \otimes I$ on the composite system. Similarly, the second formula says that the measurement of A^{II} on the system II alone gives the same result as a measurement of $I \otimes A^{\mathrm{II}}$ on the system I + II, These relations must hold for all operators A^{I} and A^{II}, in particular for projections. These conditions determine W^{I} and W^{II} uniquely. The result is expressible in the following simple formulae [22])

$$\left.\begin{aligned} W^{\mathrm{I}} &= \phi\,\phi^{\#}, \\ W^{\mathrm{II}} &= \phi^{\#}\,\phi\,. \end{aligned}\right\} \tag{11}$$

The generalization of this result to an arbitrary state W is easy. Let $W = \sum \lambda_n P_n$ be the density operator of a state in I + II. Choose in the one-dimensional range of the projections P_n an arbitrary normalized vector ϕ_n, then the states in the system I and II are represented by the density operators

$$\left.\begin{aligned} W^{\mathrm{I}} &= \sum_n \lambda_n\,\phi_n\,\phi_n^{\#}\,, \\ W^{\mathrm{II}} &= \sum_n \lambda_n\,\phi_n^{\#}\,\phi_n\,. \end{aligned}\right\} \tag{12}$$

We shall refer to formulae (11) and (12) as the *reduction formulae* and we shall call W^{I} and W^{II} the reduced states of the state W. Only (11) will be used in the following.

Let us now discuss Equation (11) a little more in detail. Consider first the case that $\phi = \varphi \otimes \psi$ where φ and ψ are both normalized and $\varphi \in \mathfrak{H}^{\mathrm{I}}$, $\psi \in \mathfrak{H}^{\mathrm{II}}$. From the definition of ϕ and $\phi^{\#}$ it follows then that

$$\phi^{\#}\, \varphi = (\varphi, \varphi)\, \psi = \psi$$

and

$$\phi\, \psi = (\psi, \psi)\, \varphi = \varphi\, .$$

Thus $\phi\, \phi^{\#}\, \varphi = \varphi$ and $\phi^{\#}\, \phi\, \psi = \psi$.

On the other hand, let φ' be orthogonal to φ, so that $(\varphi, \varphi') = 0$, then

$$\phi^{\#}\, \varphi' = (\varphi', \varphi)\, \psi = 0\, ,$$

so that $\phi\, \phi^{\#}\, \varphi' = 0$. Similarly, if ψ' is orthogonal to ψ then $\phi^{\#}\, \phi\, \psi = 0$. We see in this special case that $\phi\, \phi^{\#} = P$ is a projection in $\mathfrak{H}^{\mathrm{I}}$ with one-dimensional range and $\phi^{\#}\, \phi = Q$ is a similar projection in $\mathfrak{H}^{\mathrm{II}}$. We have thus established: if $\phi = \varphi \otimes \psi$ then the reduced states are pure. This condition is also necessary, and so we find:

> the reduced states W^{I} and W^{II} are pure if and only if the pure state ϕ is of the form $\phi = \varphi \otimes \psi$.

Let us now consider the case that ϕ is still a pure state, but not of this form. Then we know from the preceding discussion that neither W^{I} nor W^{II} can be pure. Let $W^{\mathrm{I}} = \sum \alpha_r\, P_r$ with P_r a projection of one-dimensional range and $\alpha_r > 0$, $\sum_r \alpha_r = 1$. Define

$$\psi_r = \frac{1}{\sqrt{\alpha_r}}\, \phi^{\#} \varphi_r\, .$$

It follows that

$$W^{\mathrm{II}}\, \psi_r = \frac{1}{\sqrt{\alpha_r}}\, \phi^{\#}\, \phi \phi^{\#}\, \varphi_r = \frac{1}{\sqrt{\alpha_r}}\, \phi^{\#}\, W^{\mathrm{I}}\, \varphi_r = \sqrt{\alpha_r}\, \phi^{\#}\, \varphi_r = \alpha_r\, \psi_r$$

and

$$||\, \psi_r\, ||^2 = \frac{1}{\alpha_r}\, ||\, \phi^{\#}\, \varphi_r\, ||^2 = \frac{1}{\alpha_r}\, (\varphi_r, \phi\, \phi^{\#}\, \varphi_r) = 1\, .$$

Thus ψ_r is a normalized eigenstate of W^{II} with eigenvalue α_r. Furthermore, every eigenstate is of this form. Thus the operator $\phi^{\#}$ establishes a correspondance between the eigenstates φ_r of W^{I} and the eigenstates ψ_r of W^{II}. It follows from this that W^{II} has the form $W^{\mathrm{II}} = \sum \alpha_r\, Q_r$ with $Q_r\, \psi_r = \psi_r$. Furthermore, one verifies easily that $\phi = \sum_r \sqrt{\alpha_r}\, \varphi_r \otimes \psi_r$. Thus we find: if ϕ is a general state vector in $\mathfrak{H}^{\mathrm{I}} \otimes \mathfrak{H}^{\mathrm{II}}$ then there exists an orthonormal system $\{\varphi_r\}$ in $\mathfrak{H}^{\mathrm{I}}$ and an orthonormal system $\{\psi_r\}$ in $\mathfrak{H}^{\mathrm{II}}$ and a set of positive numbers α_r such that

$$\phi = \sum_r \sqrt{\alpha_r}\, \varphi_r \otimes \psi_r\, , \quad W^{\mathrm{I}} = \sum_r \alpha_r\, P_r\, , \quad W^{\mathrm{II}} = \sum_r \alpha_r\, Q_r\, . \tag{13}$$

We can see that the example discussed at the end of Section II gives an illustration of this general theorem. For the case that $\cos\delta > 0$, $\sin\delta > 0$, we find that the correspondence $\alpha_1 = \cos^2\delta$, $\alpha_2 = \sin^2\delta$, $\varphi_1 = u_1$, $\psi_1 = u_2$, $\varphi_2 = v_1$, $\psi_2 = i\, u_2$ brings this result in agreement with the general theorem (apart for the notation of the tensor product, which is more informal and less precise in the example).

We have now all the formal apparatus assembled that is needed for the detailed discussion of the measuring process. This we shall do in the next Section.

VI. The Measuring Process

The measuring process is that physical process which permits the determination of the truth or falsehood of one or several propositions about a physical system[23]). A measuring process always requires an observation on the system. An observation means an interaction of the system S to be measured with another system M, the measuring device. The detailed analysis of the measuring process requires thus the study of two systems in interaction with one another. The measuring device must respond differently on a macroscopic scale to different states of the system to be measured. Since quantum systems are usually microscopic in nature, this means that the measuring system should comprise a metastable *amplifying device* that can be triggered by microscopic events. Examples of such systems are superheated liquids (bubble chamber), supercooled vapours (Wilson chamber), a neutral gas in a strong electric field (counters or spark chambers), etc.

This is the usual description of the measuring process and what we have to add are merely some comments in order to ensure that the notions are precise and relevant. The macroscopic nature of the measuring apparatus is of course not a precise notion. However, it is not the large size, it is the classical aspect of the measuring apparatus which is essential. It is a well-known fact that large objects usually behave classically, but there are large systems which exhibit quantum phenomena and conversely a small system with a suitably restricted set of observables may very well behave classically. The notion of classical system *can* be made precise and it means a system with an Abelian set of observable. The notion has only a meaning with respect to a specified set of observables. If we look hard enough, every system has quantum features. The erroneous conclusion that there are no classical systems is only possible if one forgets that the notion of 'classical' has only meaning with respect to a specified set of observables.

The measuring process reduces the measurable alternatives to those of a classical state which, because of its classical nature, has an objective reality in the sense of Section IV. Its invariance under observation guarantees that the measurable alternatives have become 'events' which can at any time become 'data' without in any way affecting the state.

We see from this that the amplifying device which is the usual ingredient of a measuring device is only necessary if the measured event is to become a datum for some observer. To be sure, some last step of this kind is always necessary if the measurement is to be useful as a piece of information. The point we are making here is that this last step (changing an event to a datum) is nothing which involves any specifically quantum mechanical features. It is common to all of classical physics. It is just as much involved in the observations which reveal the alternatives of a thrown die. There is therefore no reason to pay any more attention to the macroscopic observer in the quantum mechanical measuring process as in the corresponding classical process.

The characteristic features of events are that they can be established without interferring with the state of the system, as we have explained in detail in Section IV.

They have therefore an objective character in the sense that if they are recorded they will be recorded identically by all observers who have sufficiently sensitive organs to perceive the event. They become therefore also communicable and they can thus furnish the raw material of any theoretical picture of nature.

Due to this situation, it is possible to simplify and clarify the analysis of the measuring process if we divide the measuring apparatus into two subsystems $M = m + A$, where m denotes the microscopic part of the measuring device and A the amplifier. Instead of considering the $S + M$, we consider the system $S + m$, which then is further combined to the system $S + M$ in the last stage of the measurement, when the event is raised to the macroscopic level.

We shall now apply to the system $S + m$ the formalism of the preceding Section, denoting the system S to be measured as system I and the microscopic part of the measuring device m as the system II. The measuring device is then completed by joining to m the amplifier A, which permits the raising of the event happening in system m to the level of a datum.

The system I may be anything. The system II on the other hand, must have certain properties which make it suitable as a measuring device. It must interact with the system I in such a way that certain states of system I after the interaction with II leave the system II in orthogonal final states which can be amplified by the amplifier A to macroscopically distinguishable states.

It is possible to recognize two kinds of measurements[24]). In a *measurement of the first kind* an immediate repetition of the measurement yields the same result as before with probability 1. In a *measurement of the second kind* this is not the case and the interaction with the measuring device leaves the system in a new state which will not be an eigenstate of the measured quantity. Only the first kind of measurement can be used for the *preparation* of a state for which certain observables have definite and prescribed values. We shall be concerned here exclusively with measurements of the first kind.

Let us now analyze the measuring process in a model situation which still contains the essential features of the general situation, but which omits inessential details.

We take for system I a system with a two-dimensional state vector space. Let φ_+ and φ_- be two orthogonal vectors in this space which are eigenstates of the quantity to be measured. The state vectors of system I are a three-dimensional space. It contains a state ψ_0 which describes the 'state of readiness' of the measuring device and two more states, denoted by ψ_+ and ψ_-. Here ψ_+ is the final state of II after it has measured the system I in the initial state φ_+; likewise ψ_- is associated with the state φ_-. Let $\phi_{+0} \equiv \varphi_+ \otimes \psi_0$ be the pure initial state vector before I and II begin to interact. After the interaction, the joint system is in the pure state

$$\phi_+ = U\,\phi_{+0} = \varphi_+ \otimes \psi_+ \,,$$

where U is some unitary operator. Similarly $\phi_{-0} \equiv \varphi_- \otimes \psi_0$ is sent by the interaction into the state

$$\phi_- = U\,\phi_{-0} = \varphi_- \otimes \psi_- \,.$$

In these formulae we have expressed precisely (in the framework of this model) what was contained in the description of a measurement of the first kind.

Let us now consider the general initial state of the form

$$\phi_0 = \alpha_+ \phi_{+0} + \alpha_- \phi_{-0} ,$$

where $\alpha_\pm$ are two arbitrary constants subject to the normalization condition $| \alpha_+ |^2 + | \alpha_- |^2 = 1$. Since U is a linear operator, we must have for the final state after the measurement in this case

$$\phi = U \phi_0 = \alpha_+ \varphi_+ \otimes \psi_+ + \alpha_- \varphi_- \otimes \psi_- . \tag{14}$$

This is the state of system I + II after the measurement. It is pure, as it must be, since the initial state was pure. Why does this state furnish us with a measurement of the alternatives φ_+, φ_-? The reason is that, reduced to the measuring system II, this state is a mixture, therefore each individual system realizes one of the events of the mixture and these events can, by virtue of the amplifying device A, be made data for an observer. The relevant part of the state ϕ is thus its reduction to the system II. In order to carry out this reduction we must use formula (11) of Section V and we find for the reduced density operator of the two systems after a simple calculation

$$W^{\mathrm{I}} = | \alpha_+ |^2 P_+ + | \alpha_- |^2 P_- , \tag{15}$$

$$W^{\mathrm{II}} = | \alpha_+ |^2 Q_+ + | \alpha_- |^2 Q_- , \tag{16}$$

where we have introduced the projection operators $P_\pm$, $Q_\pm$ with one-dimensional ranges containing $\varphi_\pm$, $\psi_\pm$ respectively.

This result shows clearly the reason for the occurrence of process (1) under the influence of a measurement. It is here merely a mathematical consequence of the reduction of a pure state (14) to one of its component subspaces.

The problem poses itself differently in case we combine the measuring device M with the system S to a larger (classical) system $S + M$ which is now simultaneously a system and a measuring apparatus. This is done for instance in the discussion by WIGNER[8]). In this case there is no question of reducing the state (14) to one of the component systems. There is therefore no possibility of changing the state into a statistical mixture in this manner. Yet if $S + M$ is a measuring device, it will record a mixture. This is implied in the objective character of the measurable alternatives as it was explained in Sections III and IV.

This paradoxical situation is the source of the far reaching suggestion of WIGNER to renounce the linear laws of quantum mechanics in the description of the measuring process[9]). We shall show now, again within the limitations of this model situation, that there is no need to do this if full account is taken of the essentially classical feature of the measuring device $S + M$.

Let us introduce the projection operators $\Pi_+ = P_+ \otimes Q_+$ and $\Pi_- = P_- \otimes Q_-$. Quantum mechanics, in accordance with process (2), tells us that the final state of the system $S + M$ is the pure state (14). However, process (1) applied to $S + M$ as a measuring instrument tells us that this final state should be

$$W = | \alpha_+ |^2 \Pi_+ + | \alpha_- |^2 \Pi_- . \tag{17}$$

Which one of these two descriptions is correct? LONDON and BAUER[6]) would say: before the datum entered consciousness it is the state (14), after consciousness has taken cognizance of the datum it is the state (17). This peculiar interpretation is certainly not in agreement with modern scientific methodology, and in this form it is one of the weakest points of quantum mechanics.

However, we maintain that it is not necessary. The state after the interaction is neither (14) nor (17), or it is both if one prefers. What this loose and picturesque language means shall now be made precise.

If $S + M$ is a measuring device, then it is a classical system. In fact, the non-trivial projection operators which are observable are the two projections Π_+ and Π_-. They commute and therefore every observable commutes with every other one. The system of observables is Abelian. Therefore according to the theory of Section IV the states of the system $S + M$ can be placed into equivalence classes.

The meaning of the above paradoxical statement is now simply this. The two states (14) and (17) are in the same equivalence class. Let us verify this.

The most general observable of $S + M$ has the form $A = \mu_+ \Pi_+ + \mu_- \Pi_-$. We must show that for any such observable

$$(\phi, A \phi) = Tr\, W A ,$$

where ϕ is the state vector (14) and W is the expression (17). A simple calculation shows that both sides of this equation are equal to $\mu_+ |\alpha_+|^2 + \mu_- |\alpha_-|^2$.

Thus, if the system $S + M$ is truly classical, the two states (14) and (17) cannot be distinguished from one another, and one of the most vexing problems of quantum mechanics dissolves into a pseudoproblem.

A final question: does this result mean that the two states (14) and (17) can, under all circumstances, never be distinguished by a measurement? It does not mean this. It means, this can never be accomplished with measurements from the Abelian set $\mathfrak{S}$ which contains the two projections Π_+ and Π_-. In order to distinguish them it is necessary to have at one's disposal an observable which is not in this set. An observation of such a quantity will no doubt reveal that it is indeed (14) which is the final state after the interaction, in agreement with the Schrödinger equation.

This conclusion does not invalidate the statement[25]) that the measurement produces the state (17) [or rather the equivalence class containing (17)] because, as BOHR has always emphasized, the very possibility of measurement implies a classical apparatus[26]).

VII. Concluding Remarks

The analysis of the measuring process presented in this paper shows that there is no evidence to question the applicability of quantum mechanics to large systems or even systems with consciousness. The attribute of an 'element of reality' to the measurable alternatives in a mixture is justified because such quantities can be observed without in any way interferring with the state of the system. This is conform to the 'classical' interpretation of states. Such states can sometimes be prepared with a maximal set of relevant conditions on the microscopic level and they furnish the indispensable classical aspect of any measuring process. This interpretation permits the omission of VON NEUMANN's 'ultimate observer' and the ultimate occurrence of process (1) during the measuring process. It also shows that the physical comprehension of the measuring process does not require a complete theory of the approximate classical behaviour of large quantum systems.

In order to assert this result in full generality, the analysis carried through in this paper for a special case should be generalized and refined in several directions. One should extend it to an observable with more than two values, and one should also allow the possibility of a degenerate spectrum for the observed quantity. Then one should take into account that both S and M could be in a mixture before the measurement begins. Furthermore, one should also include the case of continuous spectra. Finally, the discussion should then be extended to measurements of the second kind. Only then could we make these assertions in full generality. We have not done this here, since a treatment of such generality would require a considerable mathematical apparatus which would tend to obscure the essential point of our discussion. But there is no doubt that the conclusions reached here are applicable for systems of much greater generality than the ones we have discussed.

Acknowledgements

The author gratefully acknowledges an illuminating correspondence with Professor WIGNER. The critical comments by Dr. J. S. BELL of CERN have been essential in shaping th author's views expressed in this paper.

References

1) W. HEISENBERG, Z. Phys. *43*, 172 (1927).
2) L. ROSENFELD, Nature *190*, 384 (1961), gives a recent account of Bohr's point of view concerning complementarity.
3) M. BORN, Z. Phys. *38*, 803 (1926).
4) The best exposition of these controversial aspects of quantum mechanics known to me is the article by N. BOHR and EINSTEIN's *Reply to criticism*, in A. EINSTEIN, *Philosopher-Scientist*, (Library of Living Philosophers, Evanston 1949).
5) J. VON NEUMANN, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin 1932), esp. p. 186 ff.
6) An excellent presentation of the entire problem and in particular of this last point is found in the paper by F. LONDON and E. BAUER, *La Théorie de l'Observation en Mécanique Quantique*, Actualités scientifiques et industrielles, No. 775 (Hermann, Paris 1939).
7) A representative selection of some of the more recent papers is: I. FEYERABEND, Z. Phys. *148*, 551 (1957); G. SÜSSMANN, Bayer. Akad. München *88* (1958); H. MARGENAU and R. N. HILL, Prog. Theor. Phys. *26*, 727 (1961); H. WAKITA, Prog. Theor. Phys. *23*, 32 (1960); *27*, 139 (1962); A. DANIERI, A. LOINGER and G. M. PROSPERI, Nuclear Phys. *33*, 297 (1962).
8) E. P. WIGNER, Am. J. Phys. *31*, 6 (1963).
9) We cite the following passage from Reference [8]: 'A full description of a measurement remains impossible, since the quantum-mechanical equations of motion are causal and contain no statistical element, whereas the measurement does.'
10) E. P. WIGNER, *Remarks on the Mind-Body Question*, in The Scientist Speculates, edited by J. GOOD (W. Heinemann, London 1962), p. 284.
11) E. SCHRÖDINGER, Naturwiss. *23*, 807 (1935).
12) A. EINSTEIN, B. PODOLSKY and N. ROSEN, Phys. Rev. *47*, 77 (1935).
13) N. BOHR, Phys. Rev. *48*, 696 (1935).

[14]) This description of the state is of course the usual one of ordinary quantum mechanics. It is, however, possible to conceive more general possibilities of states, where for instance the probability distribution of an individual observable may depend on the exact physical conditions of the measuring device. In such cases, the basic notions would be joint probabilities for complete sets of compatible observables. We shall not consider this possibility here, because there seems to be no need for this generalization.

[15]) G. BIRKHOFF and J. VON NEUMANN, Ann. Math. *37*, 823 (1936).

[16]) C. PIRON, *Thesis*, University of Lausanne, to be published.

[17]) A. M. GLEASON, J. Math. Mech. *6*, *885* (1957).

[18]) This possibility was considered by VON NEUMANN himself, but it has not led to any tangible results ,cf. J. M. JAUCH, *Continuous Geometries and Superselection Rules*, CERN 61–14. GLEASON's theorem has only been demonstrated for the projections in a VON NEUMANN factor of type I, its generalization to factors of type II or III is an unsolved problem.

[19]) The relevance of the theorem of VON NEUMANN, alluded to here, has been questioned by some authors. The measuring problem is discussed from this point of view for instance in L. DE BROGLIE, *La Théorie de la Mesure en Mécanique Quanitque* (Gauthier-Villars 1957) and D. BOHM, *Causality and Chance in Modern Physics* (Routledge and Kegan Paul, London 1958), where further references can be found. A recent re-examination of this theorem has shown that VON NEUMANN's demonstration can be generalized so that most of the objections against it can be devaluated. This question is treated in a separate publication by J. M. JAUCH and C. PIRON, Helv. Phys. Acta *36*, 827 (1963). The only possibility which remains perhaps is to generalize the notion of 'state' in the sense of footnote [14]) which of course would mean a radical departure from quantum mechanics with consequences which cannot yet be foreseen.

[20]) The notions which we have introduced here are closely related to the contents of a paper by W. FURRY, Phys. Rev. *49*, 393 (1936), in which he criticizes the paper by EINSTEIN, PODOLSKY and ROSEN (cf. ref. [12])). In fact, what FURRY shows (in our terminology) is that the states to which EINSTEIN *et al.* attribute an 'element of reality' are in fact not classical states under all imaginable observations.

[21]) D. FINKELSTEIN, J. M. JAUCH, S. SCHIMINOVICH and D. SPEISER, J. Math. Phys. *3*, 207 (1962), esp. Section 4.

[22]) Cf., Ref. [5]), p. 230.

[23]) We use the word 'proposition' in the technical sense here to designate yes-no experiments. Every measurement of a physical quantity can be broken up into the measurement of a suitably chosen set of such propositions.

[24]) W. PAULI, *Handbuch der Physik*, Vol. 1, p. 72 ff.

[25]) The tought experiment constructed by FURRY (Ref. [20])) may be considered as a demonstration of this assertion.

[26]) The similarity of this situation with the classical notion of temperature in the presence of Maxwell demons is striking. I refer here to a remark by CLAUSIUS which is quoted in the same context by ROSENFELD in Ref. [2]).