Foundations of Quantum Mechanics

Jauch, Joseph-Maria

Foundations of Quantum Mechanics

Lectures delivered at the Varenna School

summer 1970

by

J. M. JAUCH, Geneva Switzerland

UNIVERSITÉ DE GENÈVE

INSTITUT DE PHYSIQUE THÉORIQUE
ÉCOLE DE PHYSIQUE, CH-1211 GENÈVE 4

I. The conceptual structure of quantum mechanics

1) Quantum Mechanics, as we know it today, was born out of the need to explain a whole series of observed facts in atomic and molecular physics, which could not be reconciled with the concepts of classical physics. Among them were the laws of black body radiation, the specific heat of solids, the structure of the atoms and their interaction with radiation, the photoelectric and the Comptoneffect and many others.

The various stages of development, beginning with Plancks bold and singularly successful hypothesis of the quantization of the radiation field oscillators, passing through the older form of the quantum conditions for periodic motions culminated in the almost simultaneous discovery of matrix-mechanics by Heisenberg and wave-mechanics by Schrödinger. It was soon recognized that these apparantly so different theories were two different aspects of one all-embracing theory, now called quantum mechanics.

Dirac succeeded in formalizing the structural analogy between quantum mechanics and classical mechanics by exhibiting the correspondence between classical Poisson brackets and quantum commutators, thus incorporating into the theory with one ingenious device the older correspondence principle by Bohr.

Yet, the formal analogy between classical and quantum mechanics was deceptive. There seemed to be a profound gulf which separated the two theories. It was brought to light in the now generally accepted probabilistic interpretation of the wave function, the uncertainty relations, or the more general concept of complementarity introduced by Bohr. It is a most remarkable fact, with no parallel in the history of science, that although quantum mechanics has become the indispensable basic theory for all of microphysics and for much of macrophysics as well, its interpretation has remained a source of conflict from its inception in the late twenties until today. The epistemological questions associated with the so called Copenhagen interpretation are so profound and revolutionary that many of the founders of this theory (among them Einstein, de Broglie and Schrödinger) have rejected this interpretation and for many thoughtful physicists of a more practical bent of mind it has remained a kind of "sceleton in the clauset".

2) There have been numerous expositions of these difficulties, some in the form of paradoxes, as for instance the paradox of Einstein, Podolsky and Rosen[13)], the "Schrödinger cat" or "Wigner's friend", others consisting of a more formal analysis of the measuring process or the philosophical or mathematical discussions on hidden variables. The fact that these expositions often take the form of controversies with ideological overtones and that they have been going on for more than forty years and have even lately shown signs of increasing in volume and intensity is perhaps sufficient evidence that quantum mechanics has left us with an enigma notwithstanding the profound work of Bohr and the "Copenhagen school".

It is perhaps rather strange that none of the various proposals for a more satisfactory reinterpretation of the quantum mechanical formalism have led to the prediction of any observable consequences, nor has quantum mechanics so far been generalized in any physically significant manner. Attempts in this direction are numerous but none can be really considered as being successful. Thus we shall have to deal with the quantum mechanics as it was created in the late twenties a theory which has been tested by an immense variety of the most refined experiments of which none has given the slightest indication that anything might be wrong with it. We should of course not confuse the issue by pointing to the considerable number of unexplained facts such as the existence of elementary particles (or anything for that matter) their various types of interactions and their symmetry properties. All these things are outside the predictive power of quantum mechanics, but they do not contradict it. Indeed every one can be incorporated into the conceptual framework of the theory and there is absolutely no indication that quantum mechanics might not be the fundamental theory sufficiently powerful to embrace all the facts known today of the microsystems.

This situation is entirely different from the one before the discovery of quantum mechanics. Then an ever increasing body of experimental facts were being accumulated which definitely contradicted classical physics and only the discovery of quantum mechanics enabled us to understand and predict such facts completely outside the scope of classical mechanics. Today there are no facts known which contradict any conclusions from quantum mechanics. Therefore one may well ask: why this concern about the foundations of quantum mechanics ?

The answer is probably that quantum mechanics has produced a revolution in the epistemology of physics that we have not yet been able to appreciate in all its implications. Almost all the conceptual structure which constituted the background of classical physics and hence the glory of three hundred years of science were affected by the advent of quantum mechanics. Such familiar concepts as the physical system, its state, its description in space and time, its evolution, causality and determinism, are all put in question or have acquired a new meaning.

Most of all, as has been particulary emphasized by Einstein it is the notion of "objective reality" of the physical world, the cornerstone of all of classical physics, which has become pratically obliterated in the new physics. Although it is not clear what exactly is meant by "physical reality" it is conceivable that the underlying idea of this term could be made precise by physics. The analysis of Einstein, Podolsky and Rosen together with Bohr's reply has made it clear that quantum mechanics is precisely not capable of doing that and according to ones point of view one may find comfort or a source of concern in that situation.

These two different attitudes are more or less correlated with the two philosophical attitudes known as Positivism and Realism (in the modern sense of the term). On the whole, and in spite of pretensions to the contrary, physicists are usually more inclined to Realism than to Positivism. If we were positivists such a school as we are attending now would presumably make not much sense. Yet I would like to maintain that there are strictly scientific questions connected with the foundations of quantum mechanics which need not be related to any particular philosophical school of thought. It is in this sense that I would like to present the problems to which I shall address myself now.

3) I shall begin by making some preliminary remarks on Axiomatics. Over the years a considerable effort has been expended by various people to reconstruct quantum mechanics on an axiomatic basis. Since such attemps have sometimes been described as being not very useful in the elucidation of the really deep difficulties of a theory I shall first say a few words in defense of axiomatics.

In order to avoid any misconceptions one should remember that axiomatics in pure mathematics and in physics are not quite the same thing. The axiomatic of pure mathematics is essentially a formal scheme which must only satisfy the criterion of consistency and completeness. There is no need to interpret the axioms or the undefined terms (although it must be admitted many axiomatic systems are motivated by such an interpretations). In physics the interpretation of the axioms and the undefined terms is an essential part in the process. Indeed these interpretations are the link of the formal structure (the theory properly speaking) with the observed facts. The axioms therefore can not be freely chosen, subject only to completeness and consistency, but they should be inductively underpinned by a considerable body of empirical data. Without such a support on the solid ground of experience physical axiomatics is sterile. The choice among possible logically equivalent axiom systems is therefore an important task of the physicist since some may be more satisfactory from the point of view of interpretation and justification than others.

Next it should be observed that the purpose of axiomatics in physics may be formulated as three-fold:

(i) It may be considered as a formal rendering of what theoretical physicist are actually doing.

(ii) It puts in evidence the empirical foundations and serves to eliminate irrelevant elements in the theory.

(iii) It gives a solid base for rigorous mathematical deductions to be made within the theory.

A few words of explanation to each of these three points may be useful. Physicists very often do not know what they are actually doing and equally often, what they do is different from what they say they are doing. It is perhaps quite satisfactory to exhibit the mysteries of quantum mechanics in straight forward and immediately obvious common sense notions as Bohr has done so remarkably well, but straightforward and common sense notions have often ambiguous meanings, with the result that in spite of the apparent simplicity of such explanations, they may and have produced more confusion than would have been necessary.

As to the second point it is perhaps well to remember that physical theories are usually confirmed or falsified by their consequences. When a theory is to be generalized or modified a postulational formulation is particularly useful since the empirical justification can be explicitly identified in a few well defined places and possible modifications can be more easily studied and surveyed.

Finally, mathematical rigor although not always necessary and sometimes even harmful (if it obscures the physical content for instance) is nothing to be rejected if it can be obtained with relative ease, as in a well-formulated axiomatic physical theory.

These obvious advantages of axiomatics in a physical theory are tempered by a certain danger, which should not be overlooked. It is in the nature of the method that the basic concepts of an axiomatic have a precise meaning and a well-defined range of validity. This range need not always coincide with the range of validity of the actual application of this concept at the physical reality. This point should always be kept in mind especially in the discussion of the so-called paradoxes of quantum mechanics.

4) Even with these requirements on the choice of the axioms the axiomatics of a theory such as quantum mechanics may contain still much arbitrariness. One of the earliest attemps of an axiomatic of quantum mechanics goes back to Jordan, Wigner and von Neumann[1)2)] and it was adopted and further developed by Segal[3)]. The basic idea is to introduce axiomatically an algebraic structure for the algebra of all bounded observables. The original motivation for this axiomatic was of course closely related to the physical interpretation of these axioms as observables. However there were always difficulties associated with this motivation, stemming from the fact that additivity of non-compatible observables cannot be easily interpreted in physical terms. Furthermore in the later stages of the theory the algebra was changed from a Jordan-algebra with a non associative product to a C*-algebra with an associative product but non-Hermitian elements (which cannot be observables). Today Segal does no longer maintain that the elements of the algebra have anything to do with observables[4)].

Another line of approach proved more satisfactory from the point of view of interpretation. It is the one initiated by Birkhoff and von Neumann[5)] adopted and developed in different varieties by Mackey,Ludwig, Pool and Piron.

The basic structure in this system is the set of elementary propositions, interpreted as the yes-no experiments (called questions by Mackey and decision effects by Ludwig) together with the set of all states. These concepts are relatively easily interpreted in physical terms and therefore a good motivation for the axiomatics is available. However this advantage is obtained at the price of a weaker structure than an algebra. One obtains in this manner only a lattice and not an algebra. There is no motivation for an operation of addition. Consequently this operation is absent in the basic axioms. The most interesting question is therefore, how this more general structure is related to conventional quantum mechanics in Hilbertspace where addition of observables is a natural and easily defined operation. There is obviously a gap between what physicists actually do and what can be based on empirical inductively generalized evidence. Mackey[6] has bridged this gap by his Axiom VII which simply introduces the complex Hilbertspace ad hoc because "it works", this means it formalizes what physicists have always done.

Mackey indicates several possibilities of more general systems and he suggests that it would be interesting to have a thorough study of the consequences of modifying axiom VII.

A study in this direction has been made by C. Piron[7] who supplied additional axioms from which something very close to VII can be deduced. The additional axioms which are needed fall into two categories. One (called henceforth axiom (P)) concerns the weakening of the distributive law beyond the modular law (adopted by Birkhoff and von Neumann) and the other refers to the property of atomicity and the covering law. While the first has a satisfactory physical interpretation the second has resisted such an interpretation until very recently[8].

Several independent and apparently equivalent forms of the axiom (P) have been given in the literature[9][10][11][12].

We shall here exposit briefly and in a rather condensed manner a somewhat different approach to the axiomatic of quantum mechanics, incorporating the advantages of the Birkhoff-von Neumann-Piron system supplemented by a new characterization of the notion of "state" of an individual system.

In this new conception of "state" we avoid the notion of probability which was necessary in the older conception. If one introduces probabilities at this stage of the axiomatics one has difficulties of avoiding the criticism of Einstein that a state is not the attribute of an individual system but merely the statistical property of a homogeneous ensemble of similarly prepared identical systems, thus leaving open the hope that a fuller description of the individual system might be possible in a more complete theory than present-day quantum mechanics.

An additional advantage of this new concept of "state" stems from the fact that with it a physical interpretation of the second category of axioms (i.e. atomicity and the covering law) becomes possible and thereby a better physical foundation of these axioms as well as their mathematical consequences is obtained.

It is not yet certain whether this new concept of state leads to some observable consequences. The analysis of interacting quantal systems which we shall examine at the end of this chapter will show that quantum mechanics as we know it today does not give any answer concerning the evolution of the state of one of the subsystems. A more complete theory concerning this point is logically possible but not yet developed.

5) The axiomatics of Birkhoff, von Neumann and Piron is based on the primitive notions of the yes-no experiments for a system. The underlying idea is that the measurement of a sufficiently large number of such experiments would be suitable for the determination of the general physical properties of the system as well as the determination of the state of the system after a suitable preparation.

The yes-no experiment will be denoted by Greek letters α, β, ,.... . The outcome of an experiment α will be one of the two values yes (1) or no (o).

If α is a yes-no experiment, then there exists another one α^{ν}, measured with the same physical equipment, and such that if the outcome of the measurement of α is 'yes' then it is 'no' for α^{ν} and vice versa. Evidently $(\alpha^{\nu})^{\nu} = \alpha$.

Any yes-no experiment can be completely characterized by giving a description of the equipment to be used and a prescription how to execute the measurement.

With this rule we can generate new experiments from a family of given ones in the following manner. Let α_i ($i \in I$, some index set) be any family of such experiments, then we can define a new one, denoted by $\prod \alpha_i$ by the following procedure: The equipment for measuring $\prod \alpha_i$ consists of the equipment used for all of the experiments α_i. The prescription for executing the measurement $\prod \alpha_i$ consists in choosing any one of the equipments α_i <u>at random</u> and recording the outcome of its measurement as the outcome of $\prod \alpha_i$.

It follows then very easily that

$$\left(\prod \alpha_i\right)^{\nu} = \prod \alpha_i^{\nu} \tag{1}$$

The proof consists in verifying that the measurement of the two experiments always gives the same result with the above prescription.

The yes-no experiments satisfy an equivalence relation which may be defined as follows: We shall say a yes-no experiment is 'true' in a particular state of the system if its measurement will give the result yes with certainty (probability one). It is a fact that certain pairs of yes-no experiments stand with respect to each other in the relation that whenever α is 'true' then β is 'true' too. We denote this relation by $\alpha < \beta$, and verify easily that it satisfies the axioms for a partial ordering, viz

$$\begin{aligned}&(1)\quad \alpha < \alpha\\&(2)\quad \alpha < \beta \text{ and } \beta < \gamma \Rightarrow \alpha < \gamma\end{aligned} \tag{2}$$

If two yes-no experiments α_1 and α_2 satisfy the relations $\alpha_1 < \alpha_2$ and $\alpha_2 < \alpha_1$ then we shall say that they are equivalent and we write for it $\alpha_1 \sim \alpha_2$. This is an equivalence relation, that is, it satisfies

$$\begin{aligned}&(1)\quad \alpha \sim \alpha\\&(2)\quad \alpha \sim \beta \Rightarrow \beta \sim \alpha\\&(3)\quad \alpha \sim \beta, \text{ and } \beta \sim \gamma \Rightarrow \alpha \sim \gamma\end{aligned} \tag{3}$$

Let α be any yes-no experiment. The set of all yes-no experiments which are equivalent to it will be denoted by $a = \{\alpha\}$ and we call it a <u>proposition</u>. Thus more explicitly in a formula

$$a = \left\{ \alpha_i \;\middle|\; \alpha_i \sim \alpha \right\} = \left\{ \alpha \right\} \tag{4}$$

It is easy to verify that if α is 'true' in the above sense then any $\alpha_i \sim \alpha$ is also 'true'. Hence we can say the proposition a is 'true' if and only if any and hence all of its yes-no experiments are 'true'. If the proposition a is 'true' we shall call it a <u>property</u> of the system. We write $a \subset b$ if for all $\alpha \in a$ and $\beta \in b$, $\alpha < \beta$. If α_i ($i \in I$, some index set) is a family of yes-no experiments, all of which are true, then $\prod \alpha_i$ is true too. We denote by $\bigcap a_i$ the equivalence class $\{\prod \alpha_i\}$ which contains the yes-no experiment $\prod \alpha_i$.

It depends only on the equivalence classes a_i and it has the properties of the greatest lower bound, that is

$$b \subset \bigcap a_i \iff b \subset a_i \quad \forall i \in I \tag{5}$$

We can define a least upper bound by setting

$$\bigcup a_i \equiv \bigcap_{a_i \subset x} x \tag{6}$$

and we can verify that it satisfies

$$\bigcup a_i \subset b \iff a_i \subset b \quad \forall i \in I \tag{7}.$$

In particular, if $\mathcal{L}$ is the set of all propositions there exists a zero and a unit element (also called the absurd and the trivial proposition) defined by

$$\phi = \bigcap_{x \in \mathcal{L}} x \quad , \quad I = \bigcup_{x \in \mathcal{L}} x \quad . \tag{8}$$

We have thus established

The set of all propositions of a physical system is a complete lattice

Everything that has been said so far could have been applied to any physical system, classical or quantal. For classical mechanical systems with a phase space we can identify the propositions with the measurable subsets of phase-space, the ordering is set-inclusion, the greatest lower bound is the set intersection and the least upper bound the set union. It follows that such a proposition system is represented by a Boolean lattice, characterized by the distributive law

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$$
$$\text{and} \quad a \cup (b \cap c) = (a \cup b) \cap (a \cup c) \tag{9}$$

The lattices which represent quantal system do not satisfy the distributive law. This has a profound effect on the existence of the so called complements. A complement b of a proposition a is defined by the properties

$$a \cap b = \phi \quad \text{and} \quad a \cup b = I \tag{10}$$

It is a known theorem for Boolean lattices that the complement, if it exists, is unique. For non-Boolean lattices on the other hand there may exist several complements to a given element. It is useful to single out from all the possible complements a particular one which is distinguished from all the other ones by a physical property which we shall call compatibility. Thus we define: The complement b of a is a compatible complement if there exists a yes-no experiment $\alpha \in a$ such that $\alpha^{\vee} \in b$.

We denote the compatible complement with a'. All the physical systems known today satisfy

Axiom C: For every proposition $a \in \mathcal{L}$ there exists at least one compatible complement $a' \in \mathcal{L}$.

Furthermore Piron has shown[7] that all known systems satisfy also

Axiom P: If $a \subset b$ then the sublattice generated by (a,b,a',b') is Boolean.

With these two axioms one can then prove that the correspondence $a \mapsto a'$ is unique and it is an orthocomplementation, that is

$$a \subset b \quad \Rightarrow \quad b' \subset a' \qquad (11)$$
$$(a')' = a.$$

Furthermore the lattice $\mathcal{L}$ is weakly modular, so that

$$a \subset b \quad \Rightarrow \quad a \cup (a' \cap b) = b$$

6) Proceeding now to the notion of state, we begin by some heuristic remarks which tend to show that this notion has some general aspects which are completely analogous in classical and quantal physics. It is the formalisation of these general aspects which furnishes us the axiomatic structure of the states.

If one analyses the statements made by physicists about the epistemology of physics one finds that implicitly or explicitly these statements presuppose a bias in favor of philosophical realism. This is perhaps naiv and it takes a considerable amount of philosophical sophistication to doubt a reasonable statement as it is for instance expressed by B. d'Espagnat[14] as

Hypothesis C: Physical systems exist and have well defined physical properties independent of any observation of these properties.

It is rather difficult, if not impossible to say what the exact meaning of such a hypothesis could be. The best possibility for a meaningful interpretation of such a statement seems to me to regard it as a summary of a large number of inductive generalizations based on practically continuous and everyday experiences.

We know of course that we cannot specify a priori what these physical properties may be which we can attribute to a physical system. The existence of the uncertainty relations has taught us to be cautious in this respect. At any rate the classical attributes, exemplified by the points in phase space are not suitable for this purpose.

The difficulty is even more serious than that. The analysis made by d'Espagnat of interacting quantal systems has led him to the conclusion that hypothesis C is incompatible with the behaviour of such quantal systems[14]. I do not see how one can escape this conclusion, as long as one adhears to the usual definition of state. But it seems possible to modify this definition in such a way that this conclusion does not necessarily hold.

In the new definition of state the basic undefined concept is the notion of 'true' for a proposition. We have in the previous section introduced this notion by saying that a is 'true' if and only if any and hence all of its yes-no experiments $\alpha \in a$ are 'true'. Although this definition sounds quite reasonable its practical application always encounters the characteristic difficulty that we can never be sure whether in a given state a proposition is really true or not unless we make a measurement of one of its experiments. Such an experiment may however alter the state of the system in just such a way that the experiment in question appears as true (or not true, as the case may be) quite independent of the state of the system. We thus have a viscious circle which can only be broken if we postulate in accord with the realistic interpretation of states that in every physical system in a given state there exists a certain subset of propositions which are 'true' independent of any measurements on the system. We shall furthermore postulate that the subset of all 'true' propositions characterizes the state so that we may make the following

<u>Definition</u>: A state of a system is the set S of all 'true' propositions of the system:

$$S = \left\{ x \;\middle|\; x \in \mathcal{L} \;,\; x \text{ is 'true'} \right\} \tag{12}$$

The following remarks may be useful for clarifying the meaning of this definition

(i) It is clear that this definition is equally applicable for classical and quantal systems. In the classical case propositions are subsets A of the phase space. A state is a point x_o in phase space and the 'true' propositions are represented by all those subsets A which contain the point x_o.

Thus

$$A \text{ is 'true' if } x_o \in A.$$

(ii) The absence of any probability statement makes it possible to attribute a state to an individual system and not only to a statistical ensemble. We shall accordingly assume that every individual system be it classical or quantal is always in a definite state.

(iii) The states defined here correspond to the pure states of ordinary quantum mechanics. We shall thus postulate every individual system is in a pure state. Mixtures do appear in quantum mechanics, they are represented by density matrices or by a lattice theoretic probability measure, but they are the attributes of ensembles only and not of individuals.

(iv) One should distinguish between the state of the system and the amount of information available about the state. Just as in the classical statistical mechanics we may think of the state as representing the maximal amount of information on the individual system. It is thus natural to assume that no two states S_1 and S_2 can be subsets of one another.

The following properties (except (4) are consequences of these remarks

(1) If $x \in \mathcal{S}$ and $x \subset y$ then $y \in \mathcal{S}$.

(2) If $x, y \in S$ then $x \cap y \in \mathcal{S}$.

(2') If $x_i \in S$ $(i \in I)$ then $\bigcap_{i \in I} x_i \in \mathcal{S}$.

(3) $\phi \notin S$, $I \in \mathcal{S}$ for every state.

(4) For any $x \neq \phi$ there exists at least one state S such that $x \in \mathcal{S}$

Property (4) is simply a requirement that the number of states should be sufficiently numerous.

From the above properties of states one can immediately draw the conclusion that the lattice $\mathcal{S}$ must be atomic[15]. Furthermore if there exists a certain class of measurements (or filters), which we have defined as ideal, then one can show that the lattice must satisfy the covering law[16]. These are precisely the additional axioms which are needed in order to have a sufficiently strong axiomatics to give a quantum mechanics in Hilbertspace (without specifying the nature of the field entirely; see however ref.[17]). It is satisfactory that with this notion of state the physical foundations of these axioms can be improved.

It is instructive to compare the notion of state introduced here for quantal and classical systems. The analogue of the classical phase space is the set of all atomic propositions. Just as we do in classical physics so here, too, could we introduce the notion of false propositions. Indeed a natural definition would be to say a is 'false' if a' is 'true'. But while in classical physics every proposition is either true or false so that the set of propositions not in the state s is necessarily fals there are in general many propositions in quantal systems which for a given state are neither true nor fals. The set of propositions, seen as a propositional calculus, behave like a non-Aristotelian logic.

7) States evolve in time just as they do in classical systems. But here, too, there is an important distinction between classical and quantal systems.

There are three kinds of evolutions to be considered for a quantal system

(a) Evolution of a closed system,
(b) Evolution of a system in contact with a classical system,
(c) Evolution of a system in contact with another quantal system.

Concerning the cases (a) and (b) the situation is very similar to that of classical systems. The essential property is that in both cases (a) and (b) the evolution is causal (or deterministic) in the sense that given the initial state then the state a certain time t later is entirely determined by this initial condition. In case (a) and in case (b) with time-independent coupling (or constant external forces) the evolution has even the group property just as in classical mechanics.

In case (c) we encounter something new with no analogue in classical physics. The evolution of an individual system is no longer deterministic. This can be seen, for exemple, in ordinary quantum mechanics where time evolution of one of two coupled quantal system cannot be described by a Schrödinger equation, a pure state is changed into a mixture in the course of time. Since we maintain that an individual system is always in a pure state the only way we can remain in accord with ordinary quantum mechanics (which we assume sufficiently well established to serve as a norm) is to postulate that an individual subsystem of two coupled quantal systems evolves stochastically. Of course we know (again from ordinary quantum mechanics) what the density matrix for the subsystems must be at each instance of time but as is well known, this information does not suffice to determine the probability distributions in phase space for a quantal system. (cf. Appendix II).

This point is related to the notion of mixtures of the 1st and 2d kind (discussed by d'Espagnat[14)]). The mixtures of the 1st kind are those which are obtained from an ensemble of which every individual is in a pure state. The mixture of the 2d kind are obtained as the state of a subsystem of two coupled quantal system. Here the usual rules of quantum mechanics do not give us any information as to the state of an individual system from the ensemble. Hence this kind of mixture is of a different kind than the first.

We may thus say the essence of our new notion of state is contained in the statement: Mixtures of the 2d kind do not exist.

If one recalls that the conclusion of d'Espagnat's critical examination of the compatibility of realism with quantum mechanics depends precisely on the occurence of mixtures of the 2d kind, one sees that this conclusion no longer follows if such mixtures do not exist.

The difficulty with this interpretation of states of subsystems is connected with the question that the notion of subsystem is itself partly ambiguous. Yet it is essential for the entire foundation not only of quantum mechanics itself but of all of microphysics. From the analysis of interacting quantal systems it becomes clear that this notion can only be used in some approximate sense. Quantum mechanics is based on this approximate notion, which for the purposes of microphysics has been found sufficient.

The pardoxical situations which one encounters in the study of interacting quantal systems leave us with the suspicion that the notion of partial system or subsystem is perhaps not compatible with the structure of the quantal laws.

If, on the other hand we take the view, as we have done so far, that states can be attributed to individual subsystems, then we are led to the conclusion that such states must envolve in general stochastically according to a law which is not furnished by quantum mechanics as we know it today.

Two questions arise immediately:

1. Is it possible to find such a law of stochastic evolution for individual subsystems which is compatible with the known average properties of such subsystems ?

2. Are there any observable consequences of such evolution which could in principle be used to test such an extension of quantum mechanics?

It is immediately clear that any observation which refers to an individual or even a statistical ensemble of a subsystem can never yield an experimental effect which could test such a theory. The reason is that such measurements are expressible as expectation-values of operators which depend only on the density matrix of the subsystem and that quantity is unchanged in any of the possible completions of quantum mechanics.

Thus the only possible experiments which could effectively test such a theory are corrolation measurements of quantities which refer to both systems. Under these circumstances it becomes questionable whether we are testing the state of an individual subsystem or not simply the state of the joint system. Since the two tests would in general yield different results they can be experimentally distinguished. We shall illustrate this in the following section by analysing the polarization correlation experiment for photons, an experiment which is actually feasible, and which answers the question that we have posed.

8) We shall consider the simplest example of two coupled quantal systems, viz. two systems of spin $\frac{1}{2}$, and we shall disregard the spatial coordinates of these spins. Such a system, by a slight shift of interpretation also describes the states of polarization of a photon[26][27]. Each of the

systems is represented by a complex Hilbertspace of dimension 2. We denote by u_+ and u_- the normalized eigenvectors of the 3-component of the spin for the first system and by v_+ and v_- the corresponding eigenvectors for the second system. We further suppose that the initial state of the system is a singlet state with total spin o represented by a state vector

$$\Psi = \frac{1}{\sqrt{2}} \left(u_+ v_- - u_- v_+ \right) \tag{13}$$

Its density matrix P_Ψ is then given by the projection of rank one into the subspace which contains the vector Ψ.

The states of the individual system (denoted by I and II respectively) are then given by density matrices W^I and W^{II}, where

$$W^I = Tr^{II} W , \quad W^{II} = Tr^I W \tag{14}$$

Here T_r^I and T_r^{II} represent the partial trace with respect to the variables of the system I or II. Thus for instance

$$W^I_{r_1 r_2} = \sum_s \left(u_{r_1} v_s , P_\Psi \, u_{r_2} v_s \right) = \frac{1}{2} \delta_{r_1 r_2} \tag{15}$$

which shows that W^I is $\frac{1}{2}$ times the unit operator I in the Hilbertspace $\mathcal{H}^I$ and similarly one finds that W^{II} is the corresponding unit operator II in the Hilbertspace $\mathcal{H}^{II}$;

$$W^{I} = \frac{1}{2} I$$
$$W^{II} = \frac{1}{2} II \qquad (16)$$

The measurement of any observable quantity A which refers to the state of the system I alone will yield a result which could be calculated with the density matrix W^I. In fact the expectation value $\langle A \rangle$ for A in the state W^I is given by

$$\langle A \rangle = Tr(W^I A) .$$

This result is obviously independent of the distribution of the pure states in the ensemble represented by the density matrix W^I. Thus with a measurement of this kind the different distributions can never be distinguished from each other. The same is of course true for measurements of quantities B which refer to the system II only.

The only kind of measurements which could possibly distinguish different distributions associated with the same density matrices W^I and W^{II} would be correlations between measurements of quantities in the two systems.

One such correlation measurement would be the measurement of the spin in two different directions in the two systems.

Let us introduce the following notation:

The unit vector α in $\mathbb{R}^3$ denotes a point on the unit sphere, P_α the projection in $\mathcal{H}^I$ and $Q_{-\alpha}$ the projection in $\mathcal{H}^{II}$ associated with this direction α. In particular for the positive 3-direction ($\alpha=(o,o,1)$) we write simply P and for the negative 3-direction ($\alpha=(o,o,-1)$) we write Q.

The projection operator $P\,Q$ represents the observable for the quantity "spin I is up" and "spin II is down".

The probability of observing $P\,Q$ in the state Ψ is then given by $(\Psi, P\,Q\,\Psi)$ which is easily calculated to be

$$x = (\Psi, PQ\,\Psi) = \frac{1}{2}. \tag{17}$$

The probability of observing P or Q alone on the other hand is given by

$$y = (\Psi, P\Psi) = (\Psi, Q\,\Psi) = \frac{1}{2} \tag{18}$$

The correlation between "spin I up" and "spin II down" is thus given by

$$\kappa = \frac{x}{y} = \frac{(\Psi, PQ\,\Psi)}{(\Psi, P\,\Psi)} = 1, \tag{19}$$

which means that whenever spin I is up, spin II is down with probability 1 (and vice versa of course). All this is well known and we repeat it here only because we want to compare it with other possibilities when adopting the new conception of states.

In order to see this let us first remark that the particular choice of the 3-direction for such a correlation experiment is irrelevant for the result. If we had chosen another direction of quantization, we would have obtained the same result: Spin I in one direction α is always correlated 100% with spin II in the opposite direction $-\alpha$.

Let us now analyze this same experiment with the new concept of state. Every observation of spin I up on system

I is do be calculated under the assumption that system I is in a pure state, given for example by the projection P_α. The probability of measuring the spin P for this system may thus be calculated by the usual rules of quantum mechanics:

$$\langle P \rangle_\alpha = Tr\,(P_\alpha P). \tag{20}$$

If the ensemble of W^I is given by the probability distribution $\rho(\alpha)$ on the unit sphere, then we find

$$y = \langle P \rangle = \int \rho(\alpha)\, Tr\,(P_\alpha P)\, d\alpha = Tr\,(W^I P) = \frac{1}{2} \tag{21}$$

for the probability of measuring "spin I up".

If angular momentum is conserved for every individual system then for every state of spin I in direction α there is associated a state of spin II in direction $-\alpha$. Thus the quantity designated before by x is now given by

$$x = \int \rho(\alpha)\, Tr\,(P_\alpha P)\, Tr\,(Q_\alpha Q)\, d\alpha \tag{22}$$

while y is given, as before, by (21)

$$y = Tr\,(W^I P) = (\Psi, P\Psi) = \frac{1}{2} \tag{23}$$

The quantity $\kappa = \frac{x}{y}$

will now assume various values depending on the probability distribution $\rho(\alpha)$. We give the result for three special cases:

(i) Uniform distribution $\rho(\alpha) = \frac{1}{4\pi}$

$$\kappa = \frac{2}{3} \tag{24}$$

(ii) All spins are in the 3-direction

$$\rho(\alpha) = \frac{1}{2}\,\delta(\alpha, 3) + \frac{1}{2}\,\delta(\alpha, -3)$$

in a symbolic notation, indicating that $\rho(\alpha)$ is a discrete measure concentrated at the points $\alpha = (o,o,1)$ and $\alpha = (o,o,-1)$ of the unit sphere. In this case

$$\kappa = 1 . \tag{25}$$

(iii) All spins are in the 2-direction

$$\rho(\alpha) = \frac{1}{2}\,\delta(\alpha, 2) + \frac{1}{2}\,\delta(\alpha, -2)$$

The result in this case is

$$\kappa = \frac{1}{2} \tag{26}$$

We see in the spherically symmetrie case (i) the correlation κ is < 1. The only case where it is equal to 1 is (ii) and this result is not spherically symmetrical.

Thus in any case there is discrepancy with the result (19) of ordinary quantum mechanics. The new concept of state should therefore be experimentally distinguishable from the old one.

The formula (22) for the correlation function of the two spin states bears a remarkable resemblance with the corresponding formula for this expression under the assumption of the existence of local hidden variables[28] although it must be born in mind that the interpretation of this formula is quite different. Instead of attempting to reduce the observed results for statistical measurements on quantal system to an underlying deterministic classical structure (hidden variables) the formula (22) has a purely quantal origin. The basic laws for microphysics are considered quantal and not classical.

9) The situation that we have analysed in the preceding section is actually realizable and it has been the subject of various experimental test and theoretical discussions.

The physical system consists of a positive and negative electron in a bound state, the positronium. The lowest state of this system is a 'S-state and it has a lifetime of $\tau = 1.25 \times 10^{-10}$ sec[18]. It decays into two photons whose polarizations are correlated in exactly the same manner as the spin $\frac{1}{2}$ discussed in the preceding section. One expects therefore the corresponding polarization correlations for the two photons which result from the positronium decay. The effect was predicted by J.A. Wheeler[19] in 1946 and measured by several groups of experimenters [20][21][22][29][30].

The experiments consist in the measurement of the correlation of Compton scattering of the two photons in the azimenthal plane, once perpendicular and once parallel. The theory for this effect was worked out by several people[23][24].

The ratio $R = \frac{C_\perp}{C_\parallel}$ of the counting rates in the two azimenthal planes depends on the two polar angles.

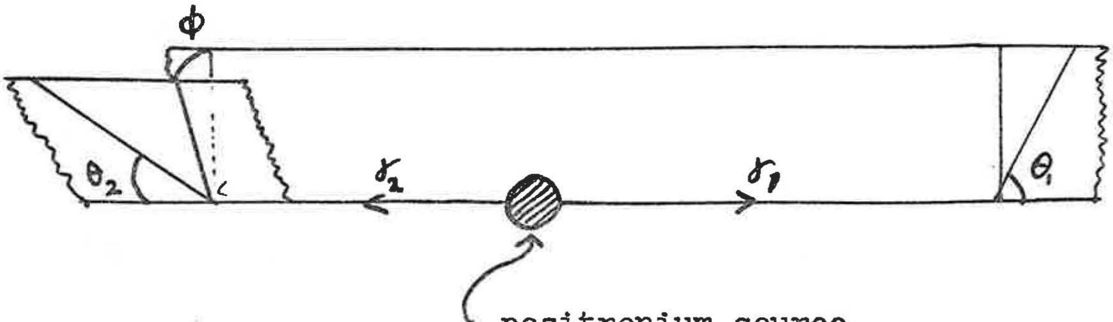

Fig.1 Azimuth (ϕ) and polar angles (θ_1 and θ_2) for the correlation measurements of Compton scattering.

The maximum value for $R = \frac{C_\perp}{C_\parallel}$ is predicted for polar angles $\theta_1 \simeq \theta_2 \simeq 82^0$ and it has a value of R = 2.8 in this case. The calculation is based on the usual formula (17). In the new definition of state, for uniform distribution of the probabilities (Equ. (24) the effect would be still observable but R would have the value R = 1.8.

The first observation by Hanna established the existence of the effect but according to this author: "Further experiments are needed, however, to determine whether the discrepancies between theoretical prediction and experimental results are merely instrumental, or due to a break-down of the theory. The deviations appear too large to be dismissed as due to statistical fluctuations, especially as the independent results using different scatterers are all consitently lower than the theoretical predictions."

The results of Hanna are in agreement with the much less accurate values of Bleuler and Bradt. The most precise measurements are those of Langhoff and of Kasday. For instance according to the published results of Langhoff R = 2.5, while the value to be expected after the geometrical corrections is 2.48, well within the statistical error. The more recent as yet unpublished results of Kasday, Wu and Ullman are in complete agreement with this.

This conclusion is in agreement with that reached by Bohm and Aharonoff[25] who had proposed a similar theory, although their motivation was different. Their conclusion was based on the less accurate experiments of Wu and Shaknov[22]. The recent improved results by Langhoff and Kasday et al. completely confirm this conclusion.

This value cannot be reconciled with the value of R = 1.8 which would be obtained for a spherical distribution of correlated spin-states (formula (22)).

It is interesting to note that this value is also incompatible with any kind of hidden variable theory of the local kind proposed by many people, most recently by Bell[28][31].

It was proved some time ago by Jauch and Piron[32][35] that such hidden variables are in disagreement with the observed structure of the quantal proposition system. This proof was free from the objections[31][34] which have been raised on physical grounds against the old proof of von Neumann[33]. Nevertheless the proof was not sufficient to give any quantitative estimates of possible discrepancies in physically observable quantities. Such a quantitative test of the theory was given by Bell[28] and Wigner[36] in the form of an inequality for the correlation of certain statistical results in a quantal system consisting of two separated parts.

It was shown by Kasday[30)] that the new results on the photon polarization correlations in the decay of positronium definitely violate these inequalities, and thus, at least in this case confirm the result by Jauch and Piron that local hidden variables are incompatible with the phenomena of microphysics.

However these same results also rule out, as we have seen, the new concept of states for the subsystem of a quantal system at least for a spherical distribution of the individual pure states. Concerning this result one can make these remarks:

(i) There is a certain ambiguity as to the physical meaning of having to do with two separated physical systems to each of which one can attribute a state. It is possible that such an identification of the individuality of a subsystem requires a certain minimum of interaction with the surrounding. One reasonable hypothesis is that the subsystem has become identified as an independent system if the perturbation that it has exerted on the surrounding has produced a state which is nearly orthogonal to the initial state. Assuming that the interaction of the photon is primarily by Comptonscattering then a rough calculation shows that this condition is satisfied when the mean angle of scattering is of the order $\alpha = \frac{1}{137}$. This is the case when the photon has traversed from about 0.1 to 1 cm of solid material. To test this theory it would be intersting to repeat the photon polarization experiment with a variable amount of scattering material interposed between the source and the place of measurement.

(ii) A further possibility of distinguishing the two photons would be to measure their polarizations at different times, which would amount to measuring their Compton scattering correlation at different distances from the source.

(iii) A more general consideration concerns the fact that the evolution of the state from a (pure) state of the joint system to a possibly correlated and stochastically evolved state of two separated subsystems cannot be described by a Schrödinger equation. In particular such an evolution is not invariant under time-reversal since a part of the correlation is lost in the process. It is thus not unlikely that the only conditions under which such an evolution can be observed is under the influence of some external interference which destroys the time reversal invariance. The possibilities discussed under (i) and (ii) are of this kind. There may be others.

It is therefore my opinion that the question concerning the interpretation of states of separated subsystems is not yet cleared up and it would be of the utmost importance to develope new experimental tests to clarify this point. Since this is the central problem concerning the foundations of quantum mechanics, with repercussions for the clarification of the so-called paradoxes as well as the measuring process, it is well worth while persuing this problem further with new and better experiments.

II. The problem of measurement in quantum mechanics

1) The problem of measurement is at the very foundation of any physical theory. It is a deep problem, but only quantum mechanics has brought to light how deep it is. The central problem in the quantum theory of measurement is this: How can one reconcile the laws of quantum mechanics with the behavior of quantal systems during the measuring process ?

At first sight it may seem strange that such a problem should occur at all. Nothing of the sort occurred in classical physics where data can be obtained by measurements on physical systems in complete accord with the (classical) laws of physics. In fact whenever a measurement disturbs the system its effect can at least in principle be calculated by these very laws of physics.

Not so for quantal systems. Measurements on such systems constitute always an interference which leaves the system in a new state after the measurement. But unlike the classical situation this new state can not be calculated in the same manner as one calculates the change of state for an unperturbed system. The quantal laws will give us only a probabilistic statement for such perturbations. This leads to the famous enigma known as the "collaps of the state vector". Let us begin by illustrating this phenomenon on some simple examples:

(i) The spherical decay of a radioactive atom.

Let us imagine an isolated radioactive nucleus which can decay and emit for instance an α-particle. Let us assume further that the nucleus is sufficiently heavy so that it may be considered at rest in the center of a heavy sphere. The sphere is supposed to be coated with a sensitive film which can detect the passage of the α particle. We suppose further that the nucleus is spherically symmetrical so that the emission is decribed by a spherically symmetrical wave function $\Psi(t)$ depending on time t.

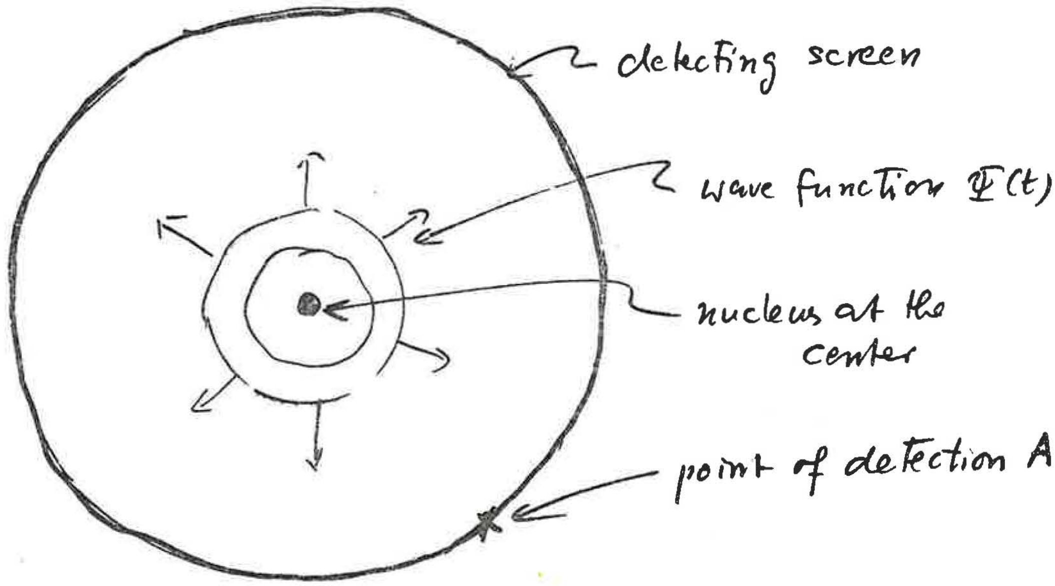

Fig.1. Detection of a decaying atom

The film on the surface is developed after the decay has occurred and there will be one point A where the α-particle has excited a sensitive grain in the film.

The difficulty in this situation is due to the fact that quantum dynamics seems unable to account for the way the spherical wave $\Psi(t)$, representing the state of the α-particle at time t, has disappeared and has been replaced by an excited state of one of the grains in the film together with an entirely different wave function concentrated near the excited grain.

(ii) Scattering of neutrons on magnetic ions.

Suppose a system of magnetic ions in a crystal lattice are all polarized in the same direction. Such a system can scatter neutrons in two different ways. There is a coherent scattering which leaves the spins of the ions as well as the neutrons unchanged. There is also a spin-flip scattering possible. In this case a neutron exchanges its spin with one of the ions. This scattering is incoherent. This can be seen as follows:

Suppose the wave function for the incoming polarized neutrons is $\varphi_0 u_+$ where φ_0 is the unperturbed spatial part of the wave function and u_+ the spin vector referring to some definite state of polarization. Let ψ_r be the scattered wave from the $r^{\underline{th}}$ neutrons without spin flip and ψ_r' the corresponding wave with spin flip, ϕ the total wave of the scatterer, and ϕ_r the total wave of the scatterer with the $r^{\underline{th}}$ neutron spin flipped.

The entire wave function after the scattering is then asymptotically

$$\Psi \simeq \varphi_0 u_+ \phi + \sum_r \psi_r u_+ \phi + \sum \psi_r' u_- \phi_r$$

The scattering cross-section is proportional to

$$\left| \sum_r \psi_r u_+ \phi + \sum_r \psi_r' u_- \phi_r \right|^2 = \left| \sum_r \psi_r \right|^2 + \sum \left| \psi_r' \right|^2$$

since $(\phi_r, \phi_s) = \delta_{rs}$ and $(\phi_r, \phi) = 0$.

Thus the spin flip term gives no interference. This result is in complete accord with the fact that a spin flip constitutes essentially a determination of the position of the neutron and that such a determination is incompatible with the interference necessary for coherent scattering.

(iii) Photon interference experiment

The typical interference experiment for photons is exemplified by a Michelson interferometer which may be represented schematically as indicated in Fig.2.

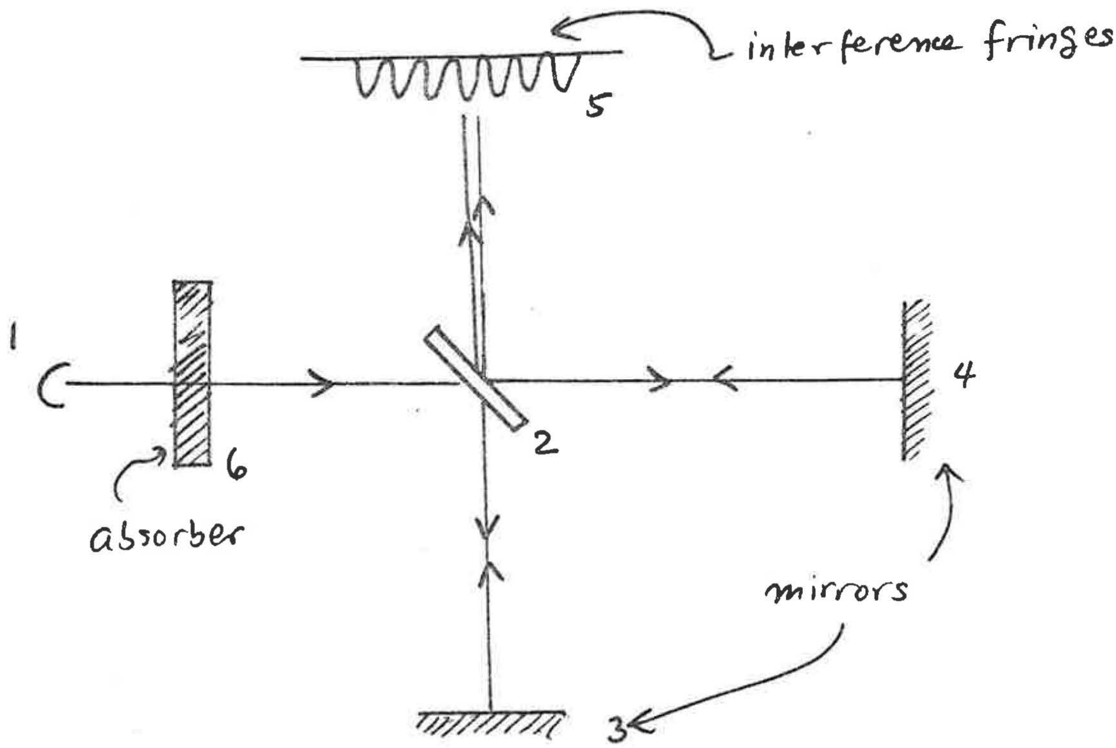

Fig.2. Interference experiment with photons

The source of light 1 sends a beam onto a half-silvered mirror which splits the beam into two, one which is reflected on the mirror 3 the other on mirror 4. The two beams are brought to interference at 5. The interference pattern at 5 can be calculated as a function of the path difference for the two beams.

This is of course a classical experiment in optics and has, in this form, nothing to do with quantum mechanics. It becomes an experiment in quantum physics if one does this experiment with individual photons. In order to do this one must reduce the intensity of the incoming light by adding an absorber 6 so that one can observe individual photons. According to quantum mechanics every individual photon should interfere with itself[37]. In order to test this point Schrödinger proposed this very experiment and it was carried out a few years ago by L. Janossy and St. Naray at the Hungarian Academy of Science.

In the final version of the experiment the intensity of the incoming light was about 10^6 photons per second. Furthermore in order to make sure that the extension of the wavepacket, which carried the photon, was smaller than the apparatus the interferometer had arms of 14 m length. This required a stability of the temperature to 10^{-5} degrees C, which was accomplished by placing the apparatus in a tunnel. The arrival of individual photons was registered by photomultiplier tubes.

Janossy verified two things : Photons arrive at the mirror positions 3 and 4 statistically independently and the classical interference pattern is reproduced to the limits of the accuracy of the experiment. Thus quantum mechanics is verified in full for this case: Photons do interfere individually with themselves.

2) The three experiment described in the preceding section are typical examples of the quantum mechanical measuring process. They concern the measurement of a property of quantal systems in a state where this property is not present or absent with certainty. This situation can only occur for quantal systems. Let us abstract from these examples the typical features of such a measuring experiment.

All of these experiments have one feature in common. They all end up with something that I shall call an event. We shall designate by event an objectively given phenomenon which has occured in a physical system irrespective whether such a phenomenon has been observed by a conscious observer. In the examples under discussion the event is the occurrence of the latent image in a silver-halide grain in experiment 1, the flipping of a neutron spin in experiment 2 and the observation of a pulse in a photomultiplier tube in experiment 3.

We shall distinguish an event from a datum. The latter is an event magnified to the level of human perception. Neither the latent image of experiment 1 nor the flipping of the spin in experiment 2 can be directly observed by a human observer and in order to make the event visible it must eventually be followed by a magnifying device which raises the event to the level of a datum. In experiment 3 the photomultiplier tube comes equipped with its own amplifier (the cascade that follows the individual absorption of a photon), so that the distinction between event and datum becomes impossible.

It is important to realize that a measurement, insofar as its quantal aspect is concerned is completed when the event has occurred. The importance of this remark lies in the fact that it helps in understanding that consciousness or other anthropomorphic concepts are in principle not involved in the measuring process. Consciousness is of course always involved in the formation of a theory but this is so in classical as well as in quantum physics. The attribution of a special rôle to consciousness in quantum physics which somehow is made responsible for the change of the statevector [38)39)] during the measuring process seems to be contradicted by the examples 1 and 2 of the previous section.

What is significent, however, is that the Schrödinger equation for the evolution of states does not describe events. This can be seen from the fact that events do not occur with certainty. For instance the silver halide grain which is affected in experiment 1 may be anywhere on the sphere and the neutron spin of experiment 2 may be anywhere inside the crystal. The final state in all these experiments are not pure states but they are mixtures.

Under the Schrödinger type evolution a pure state evolves in time t according to $\psi_t = U_t \psi$, where U_t is a one parameter continuous unitary group. Under such an evolution a pure state is always transformed into another pure state so that mixtures, as we observe then in these examples, can never be created or destroyed by such evolutions.

3) The Schrödinger evolution has another property which is incompatible with the occurrence of events. It is deterministic in the sense that the initial state determines the state at any later time with certainty. Events, however do not occur with certainty. Indeed if any one of the measurements described in the exemples of section 1 is repeated with exactly the same initial condition it would evidently be described by the same wave-function $\Psi(t)$ but it may result in a different event. In fact if a large number of such experiments under identical initial conditions are repeated one obtains a statistical distribution of such events with a probability which is given by the usual formulae of quantum mechanics. From this fact it follows again that the deterministic Schrödinger evolution does not describe events.

4) A further property which shows this fact is the irreversibility for the occurrence of events. An event, once it has occurred cannot be reversed in contrast with the Schrödinger equation which is reverisible in time. This irreversibility of the events can be seen as follows: Suppose we have followed the evolution from the time t = o to the time t = T by the mathematical solution of the Schrödinger equation and suppose further that at T an event has occurred. If we reverse the time by following this solution in the opposite direction then the state of the system Ψ_T returns to the initial state since the Schrödinger equation is invariant under time reversal. However if the system is truly reversable then the event would have to be undone and replaced by the initial state of the system. Suppose we persue the evolution once more forward in time to the point t = T. Since the initial state is the same the final state should also be the same, but this is not in agreement with the fact that under the same initial condition another event could have occurred.

Thus it is seen that all measuring processes which end up with the occurrence of an event contain an irreversible element. The very fact of the occurrence of an event precludes reversibility. The preceding argument shows that this is so precisely because the event is not determined by the initial condition. Only a certain probability distribution of various alternative events is in general determined by the initial condition.

This can for instance be verified by the experiments 1 and 2. In each of these cases the events in question (blackening of a silverhalide grain, or reversal of a neutron spin) may occur at various positions with a certain probability distribution.

The irreversibility of the quantum mechanical measuring process has often been noted before, in particular again recently where the incompatibility of this fact with the Schrödinger type evolution was emphasized[40].

There are two main possibilities of interpreting this basic aspect of the measuring process.

(i) The fundamental equations of evolution are not exactly those of quantum mechanics for systems which permit the measurement of an event.

(ii) The fundamental equations of quantum mechanics are correct for all systems, including those which are involved in a measurement, but the very process of measurement implies an approximation. This approximation in no way affects the precision of the quantities to be measured but it is precisely of the kind which wipes out the coherence of superpositions of macroscopically distinguishable states.

From a purely practical point of view there is little difference between these two points of view. Theoretically, however there is a difference since in the second version there exists at least the possibility in principle of accounting for the approximation by relating it to possible external interactions and estimating its magnitude.

In my opinion it is the second alternative which is the correct one and the detailed analysis of the examples will tend to support this point of view.

Before doing this I shall point to some rather general principles in favor of this view.

The persistence of a physical theory into regions where it has not been tested is in general a good principle to follow, be it only for the eventual determination of the limits of such a theory. Now there are in fact several large systems known which exhibit typical quantal features on a large scale. Such systems are for instance, the ferromagnet, the lasers, the superconductors. In all of these cases the unrestricted application of quantum mechanics has been successful so that we have at least some justification to assume such a validity for many more large systems.

The logical situation encountered here is similar in some respects to that of classical thermodynamics. Here, too, the irreversibility of the thermodynamics approach to equilibrium is incompatible with the reversible classical equations of motions and the former, although rigorous in a very real but limited sense, are strictly speaking only approximate as exemplified by the occurrence of fluctuations.

In quantal systems the larger the system, the smaller is the perturbation needed to wipe out the distinction between superpositions and mixtures that actually appear during a measurement process.

In fact it has been shown by Zeh[41)] that in such systems even distant interactions can have this effect so that the notion of an isolated system becomes inapplicable. The Schrödinger type evolution in such a system is then a mathematical fiction which has no real domain of validity any more.

5) In the light of the preceding discussion it is not surprising that the occurrence of events which is so characteristic for the measuring process is always associated with a certain macroscopic system for which the aforementioned approximations are valid to a very high degree of precision. In this connection the reproduction of the measuring process by a model which exhibits the necessary ergodic properties to simulate the characteristic irreversibility is of particular interest. Such models were discussed by Daneri, Prosperi and Loinger[42) and by L. Rosenfeld[43).

The discussion of such models shows quite convincingly that the measuring process is not incompatible with the fundamental equations of quantum mechanics. Furthermore the explicit use of an approximation based on the statistical properties of a large system clearly exhibits the origin of the irreversible element in the measuring process.

It should be stressed, however, that the macroscopic part of the measuring device serves two different functions which should not be confused.

The first of these functions is essential for the measuring process in the sense that it furnishes the basis for the approximation which is presupposed by the occurrence of irreversible events.

The second function is that of amplifying the event to the level of human perception, in other words, to change it into a datum. This magnifying device is almost always needed at the end of a measuring process since imperceptible events are useless as measurements. But the amplifier is merely an adapter of the event to the level of human perception and as such it is not the essential ingredient of the measuring apparatus. This is clearly illustrated by the examples 1 and 2 where the events as such are microscopic in scale and without further amplification could not be perceived by any human observer. Nevertheless we may consider the essential part of the measurement completed once the microscopic event has occurred so that the subsequent amplifier is not even mentioned in the examples. In the third example the event is the emission of a few electrons on the first plate of a photomultiplier tube, an event which would hardly be accessible to human perception. Only after the cascade is fully developed and a macroscopic pulse is registered, has the event become a datum. But again the last amplifying stage of the measuring device is not the essential ingredient of the quantum mechanical measuring device. This essential part is terminated with the absorption of a single photon on the first plate of the photomultiplier tube and as such it is a microscopic event.

This distinction of the two functions of the macroscopic measuring device is useful for separating the anthromorphic aspect of the measuring process from the physical one. In particular we emphasize again here that consciousness has no part in the essential physical part of the measuring process.

6) Let us now analyse the three examples of measurements in somewhat greater detail. The experiment 1 results in the precise location of the α-particle emitted from a centrally located source in a spherical wave.

The first point to realize is that in such an experiment the detecting grain must be localized within the desired precision of the experiment during the duraction fo the decay process. With this condition we can estimate the mass of the sphere to which the grain must be rigidly attached. Let Δx be the linear dimension of the desired region of localisation, r the radius of the sphere M its mass, so that its moment of inertia is $I = Mr^2$. Starting from the uncertainty relation

$$\Delta L \cdot \Delta\varphi \gtrsim \hbar$$

for the angular momentum uncertainty ΔL and the angle uncertainty $\Delta\varphi$, we obtain, by assuming $\Delta\varphi \lesssim \frac{\Delta x}{r}$ the inequality

$$\Delta L \gtrsim \frac{\hbar}{\Delta x} r$$

If we assume that the mean angular momentum is zero we must impose that the uncertainty ΔL is so small that the discplacement of the measuring grain during the decay time is smaller than Δx.

Thus we find $\Delta L = I \Delta\omega = Mr^2 \Delta\omega$,

$$\Delta\omega \cdot \tau \cdot r \lesssim \Delta x$$ or

$$\frac{\Delta L}{Mr} \tau \lesssim \Delta x$$

Combining this with the previously established inequality for ΔL, we find

$$\frac{\hbar\tau}{M\Delta x} << \Delta x$$

or $$M >> \frac{\hbar\tau}{(\Delta x)^2}$$

Inserting numbers for a localisation at a distance $\Delta x \sim 10^{-7}$ cm during a lifetime $\tau = 10^{3}$ sec, we find

$$M \gg 10^{-10} gr$$

Although this lower bound for the mass of the sphere is not very large compared to the usual size of macroscopic equipment, it is many orders of magnitude larger than the characteristic masses of quantal systems ($\sim 10^{-22}$-10^{-27}gr).

We may therefore still consider the "little measuring device" (in this case the grain of silver halide attached to a massive sphere) as of very small dimension although very large compared to quantal systems.

We have here a special case of a general theorem due to Wigner, Araki and Yanase concerning the size of the measuring equipment when there exists a conserved additive quantity[44)45)46)]. In this case the quantity in question is the total angular momentum, which is the sum of the angular momentum of the α-particle and that of the freely moving sphere.

A similar analysis for experiment 2 yields an inequality

$$\left(\frac{L}{\lambda}\right)^{2} \gg 1$$

where L is the linear extension of the target to which the neutron spin is rigidly attached and λ is the wave length of the neutrons which may be assumed to be of the order of the interatomic distance.

These examples show that the measuring instrument must exceed a minimal size if its quantal properties are taken into consideration and it should perform the task for which it is designed. This size is usually incompatible with the detection of phase relations between the alternatives which the measuring event selects from the various possibilities. We can see this on the example of experiment 1, for instance. The existence of measurable phase relations on the sphere presupposes that $\Delta L \lesssim \hbar$ and this inequality is incompatible with $\Delta L \geq \frac{\hbar}{\Delta x} r$ as long as

$$\frac{\Delta x}{r} \ll 1$$

As we should require if the intention of the measurement is to localize the α-particle on the sphere.

7) We may formulate the preceding consideration in a more general manner by abstracting from the examples the essential features and expressing them in a mathematical manner independent of any special features of the particular examples. The essential feature of the measuring device is that by its very construction it can admit only a limited class of observables. This means that not all self-adjoint operators in the Hilbertspace referring to the measuring equipment are in fact observables but only a very small restricted class of them. In fact the operators which are observables are in every measurement a set of commuting observables, corresponding to the fact that the measurement consists in the selection by the instrument of one of several classical alternatives.

This formulation is in complete accord with Bohr's contention that the measuring equipment must have the properties of a classical system in order to make it suitable for the establishment of objective and unambiguous events.

Bohr has put much weight on the requirement that the results of experiments, if they are to have objective validity, must be communicable in ordinary plain language perhaps suitably defined by the usual physical terminology and this requirement is guaranteed by the use of measuring instruments which allows a complete classical account of their relative positions and velocities[47)].

Many of the commentators of Bohr's fundamental analysis seem to incline to consider the crucial terms such as "classical", "objectif", and "ambiguous", which occur constantly in Bohr's writings, as undefined, or rather undefinable, as belonging to a different and prescientific epistemological level and therefore not accessible to any further analysis on the physical level.

The renunciation of any further analysis of the "classical" system introduces into the description of natural phenomena a dichotomy which cannot be exact. A classical physical system is, as we know from countless experiences, merely a limiting case of a quantal system. It is a system which for all intents and purposes behaves in accord with the laws of classical physics. But these laws, as we know, are not exact. Even the most ordinary classical object is a quantal system and its quantal aspects can be detected, provided a sufficiently refined observation is made on the system. This is true in particular for the measuring instruments that we use in the execution of a measurement on a microsystem.

It is therefore perfectly legitimate, and even necessary, to describe the interaction of the system with the measuring instrument in terms of quantum mechanics. But this description must be complemented by the specific requirements which characterizes a "classical" measuring device. In other words, if we wish to restore the unity of physical description for all physical phenomena, including measurements, we must characterize in the context of quantum mechanics the properties of a system which we want to call "classical".

According to the preceding analysis the classical system is one which has sufficiently many superselection rules that the resulting proposition system is Boolean.

In this definition "classical" appears not as an intrinsic property of the system but rather as one which expresses a limitation of the relationships of the system with the surrounding physical world. Certain propositions are excluded, not because they could not also occur under suitable conditions, but rather because they do not belong to the complete characterization of the classical system. Only in so far as the measuring system is classical, in the above sense, can it be considered a measuring device. Measurement is, in agreement with Bohr's interpretation, the result of an interaction of the microsystem with such a classical apparatus. We want to emphasize here that this conception of a measurement on quantal systems is not an arbitrarily imposed restriction but it is rather an essential ingredient in the very formalism of quantum mechanics. Indeed as it is shown in the lecture of Prof. Piron an "observable" in quantum mechanics when stripped of all nonessential elements is nothing else than a homomorphism between a classical proposition system and a subset of the quantal propositions. Such a homomorphism is mathematically the same object as a classical random variable in the theory of probability. If the quantal proposition systems are identified with the subspaces of a Hilbert space then the random variable becomes a spectral measure and via the spectral theorem it can be identified with a selfadjoint operator.

Thus when in the usual presentations of quantum mechanics an observable is represented as a selfadjoint operator then this formal property already implies that the measurement of this observable constitutes the establishment of a correspondence between the classical alternatives of a measuring instrument and a certain subset of the propositions of the system.

When we speak of classical alternatives of a measuring instrument we mean a family of mutually orthogonal projection operators Q_r in the Hilbert space $\mathcal{H}^n$ of the measuring instrument.

In the three examples cited at the beginning these projections can be identified. For instance in example (i) they represent the subspaces belonging to the excited states of the different grains of the photographic emulsion. In example (ii) they are the various orthogonal states obtained from the spin system by the reversal of one of the spins. Finally in example (iii) they correspond to the two orthogonal spaces associated with the absorption of a photon.

8) From the foregoing discussion we conclude that a typical measuring process can be described according to

the general schema indicated in Fig.3.

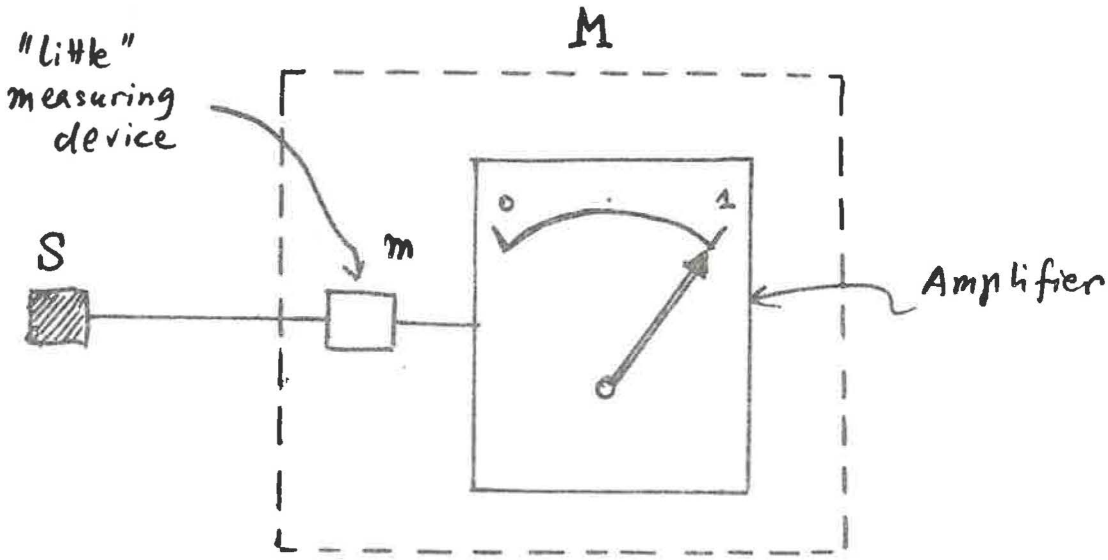

Fig.3 A typical disposition for a measurement.

The system S to be measured is in contact with a measuring device m endowed with certain properties which make it a classical measuring system. The latter is connected with the amplifier A which amplifies the event which occurs in m to the level of a datum indicated by the scale with the pointer. The little measuring device m together with the amplifier A constitutes the measuring equipment M.

The specific properties of m which make it a measuring system are first of all that it is a classical system. Thus it admits as observables only an orthogonal (and hence commuting) system of projections $Q_r (r=1,2,...)$ and all functions of this system. In general the ranges of these projections will be subspaces of some finite or infinite dimension. In order to facilitate the discussion we shall assume that they are all one-dimensional. No essential aspect of the measuring process is lost by this assumption. We then denote by ϕ_r a unit vector in the range Q_r.

We shall assume that before the measurement the system m is in a state Q_o (state vector ϕ_o) which we shall call the "state of readiness" of m. After the measurement has occurred the system m will be in another state.

For each of a certain orthogonal family $P_r (r=1,2,\cdots\ )$ of states for the system S these final states will be Q_r with certainty. If φ_r is a unit vector in the range of P_r we conclude that the unitary evolution U of the joint system S + m is such that

$$U \varphi_r \otimes \phi_o = \varphi_r' \otimes \phi_r$$

Here φ_r' represents the state of the system S after the measurement is completed.

There are two possibilities for these final states φ_r' of the system S. If we denote by P_r' the projection of range φ_r' then either $P_r' = P_r$ or $P_r' \neq P_r$. In the first case we speak of measurements of the 1st kind and in the second case of measurements of the 2^d kind. The measurements of the 1^{st} kind are also preparations of states of S of such a kind that a repetition of the measurement immediately afterwards yields the result 'yes' for P_r with certainty. The examples that we mentioned at the beginning are all measurements of the 1^{st} kind. We shall in the following assume that we are dealing with such measurements so that henceforth we use $\varphi_r' = \varphi_r$ and $P_r' = P_r$.

We obtain then

$$U \varphi_r \otimes \phi_0 = \varphi_r \otimes \phi_r \, , \tag{1}$$

For a general linear combination

$$\varphi = \sum_r \alpha_r \varphi_r \tag{2}$$

we obtain from (1) (2) and the linearity of U

$$U \varphi \otimes \phi_0 = \sum \alpha_r \varphi_r \otimes \phi_r \equiv \Psi \, . \tag{3}$$

The state Ψ is of course the state of the joint system S + m. The reduction of this state to the subsystems gives, as is well known, for each of them a density matrix which are given respectively by

$$W^S = \sum_r |\alpha_r|^2 P_r$$
$$W^m = \sum_r |\alpha_r| Q_r \tag{4}$$

According to the interpretation of states given in lecture 1 each <u>individual</u> subsystem is in a pure state but in an ensemble of such individual systems the pure states would be distributed in such a way the probabilities for all measurements are in accord with those calculated from the density matrix.

Quantum Mechanics teaches us that the pure states which may occur for the individual classical system are just precisely those which characterize the classical system, that is the projections Q_r. Thus with this assumption every individual system is in one of the states Q_r. The probability distribution which corresponds to the density operator W^m (4) is then unique since the Q_r are pairwise orthogonal and it is given by $|\alpha_r|^2$. We have thus recovered the usual rule for the probability of measuring the quantity P_r : It is given by $|\alpha_r|^2$.

9) From the foregoing discussion it is clear that the new interpretation of states of a subsystem is in complete agreement with the usual rules of quantum mechanics. Essential in this result is that the two systems S and m maintain their identity and in spite of the interaction, each remains a system of its own.

There is another interpretation possible which precludes the spearate identification of S and m. Indeed we may consider S and m together as a joint classical system Σ with the orthogonal projections $\Pi_r = P_r Q_r$ as observables.

In this case the reasoning employed before, involving the new interpretation of states is not possible and we are at a loss to explain in this manner the distinction between the pure state Ψ and the mixture $W = \sum |\alpha_r|^2 \Pi_r$

The Schrödinger evolution implies that the state after the measurement is Ψ while the actual measurement seems to say that the final state is W.

The solution of this paradox is due to the following fact: The observables of a classical system do not represent all selfadjoint operators but only a commuting subset of such operators. In the present case it is the set Π_r of orthogonal projections (and all functions thereof) which are observables. This has as a consequence that certain states cannot be distinguished from one another. They are then said to be equivalent.

More precisely if $\mathcal{O}$ is the set of all observables, we say W_1 and W_2 are equivalent with respect to $\mathcal{O}$ if

$$\mathrm{Tr}(A\, W_1) = \mathrm{Tr}(A\, W_2) \qquad \forall A \in \mathcal{O} \tag{5}$$

It is easy to see that this is an equivalence relation which we may write in the form

$$W_1 \sim W_2 \quad (\mathcal{O}) \tag{6}$$

If W is any state then we designate by $\{W\}$ the set of all states equivalent to W. An individual state W will be called a microstate and an equivalence class W a macrostate.

The initial state $\sum_r \alpha_r \varphi_r \otimes \Phi_0$ is a microstate as well as a macrostate. That is, the equivalence class which contains it contains no other microstate. But the final state is not of this kind. In fact it is easy to show that with respect to the system Π_r the two states (projection into Ψ) and W are always equivalent so that we may write

$$\Pi_\Psi \sim W \quad (\Pi_r) \tag{7}$$

Thus the two states are not distinguishable by measurements from the classical system $\overline{\Pi}_r$. There is therefore no difficulty as far as the measurements are concerned in the difference of the description of the microstates.

APPENDIX I

A mathematical model for the lattice of yes-no experiments.

The following mathematical model shows that the notion of randomness which is used in the text can be eliminated from the axiomatic. It is due to Prof. W. Reinhardt of the University of Colorado in Boulder.

We denote by $\mathcal{E}$ the set of all yes-no experiments and by Σ the set of all states.

For every $\alpha \in \mathcal{E}$ and every $s \in \Sigma$, that is every element (α, s) from the product set $\mathcal{E} \times \Sigma$, there exists a "trial set" D^s_α. This set is to some extent arbitrary but it may be identified with a countably infinite one. Let $t \in D^s_\alpha$. For any fixed $s \in \Sigma$ and $\alpha \in \mathcal{E}$ there exists a function $\alpha : D^s_\alpha \to \{o, 1\}$. We denote it by $\alpha(t)$ and refer to it as the trial outcome in the state s.

Let α, β be two elements from $\mathcal{E}$. We define an element $\alpha \bullet \beta$ by giving its domain as

$$\text{(i)} \quad D^s_{\alpha\,\beta} = D^s_\alpha \times D^s_\beta \times \{1, 2\}$$

and its values

$$\text{(ii)} \quad a \cdot \beta\,(t_\alpha, t_\beta, i) = \begin{cases} \alpha(t_\alpha) & i = 1 \\ \beta(t_\beta) & i = 2 \end{cases}$$

More generally for any family α_i $(i \in I)$ we define

$\prod \alpha_i$ by

Appendix I

(i) $D^s_{\Pi \alpha i} = \left(\bigtimes_{i \in I} D^s_{\alpha i} \right) \times I$, $t = \bigtimes_{i \in I} t_i$

(ii) $\left(\prod_{i \in I} a_i \right) (t, j) = a_j (t_j) \quad \forall j \in I, \forall t \in \bigtimes_{i \in I} D^s_{\alpha_i}$

We further define $\alpha^{\vee}$ by

(i) $D^s_{\alpha^{\vee}} = D^s_{\alpha}$

(ii) $\alpha^{\vee}(t) = 1 - \alpha(t)$.

We postulate the following two properties

(I) $a_i \in \mathcal{E} \Rightarrow \prod \alpha_i \in \mathcal{E}$

(II) $\alpha \in \mathcal{E} \Rightarrow \alpha^{\vee} \in \mathcal{E}$,

and we define

β is 'true' for the state s if and only if

$\alpha(t) = 1 \quad \forall t \in D^s_{\alpha}$

We have then the following

Theorem: $\left(\prod_{i \in I} \alpha_i \right)^{\vee} = \prod_{i \in I} \alpha_i^{\vee}$

Proof: verify

The rest of the construction proceeds now as in the text.

APPENDIX II

The problem is this : Let W be a density operator, satisfying thus

$$W^* = W > o \qquad W^2 \leq W \tag{2.1}$$
$$Tr\, W = 1$$

Is it possible to represent W in more than one way as a mixture of (not necessarily orthogonal) pure states ? The answer is obviously yes if the operator W is degenerate. For the subspace spanned by the eigenvectors of a degenerate eigenvalue is then of dimension larger than 1 and within this subspace every vector is eigenvector. Thus any orthogonal system of such vectors will give rise to a new mixture with the same density matrix. We may thus discuss the non-trivial case that all eigenvalues of W are simple. It obviously suffices to consider a two-dimensional space.

In this case we have from the spectral theorem two orthogonal projections P_1 and P_2 of rank 1 and two real numbers $\lambda_1 > o$, $\lambda_2 > o$, $\lambda_1 + \lambda_2 = 1$, $\lambda_1 \neq \lambda_2$,

such that

$$W = \lambda_1 P_1 + \lambda_2 P_2 \tag{2,2}$$

We wish to construct two other projections of rank 1, Q_1 and Q_2 and two real numbers μ_1 and μ_2 $\mu_1 > o$, $\mu_2 > o$, $\mu_1 + \mu_2 = 1$ such that

$$W = \mu_1 Q_1 + \mu_2 Q_2 \tag{2,3}$$

It is obvious that Q_1 and Q_2 cannot be orthogonal since the uniqueness of the spectral resolution of W would then imply that $Q_1 = P_1$ and $Q_2 = P_2$.

It follows from (2.3) that the projections Q_1, Q_2 must satisfy the condition

$$\mu_1 \alpha_{11} + \mu_2 \alpha_{12} = \lambda_1 \qquad (2.4)$$

$$\mu_1 \alpha_{21} + \mu_2 \alpha_{22} = \lambda_2$$

where $\alpha_{rs} = \mathrm{Tr}\,(P_r Q_s)$ (2.5)

A first condition for the possibility of a solution is thus

$$\Delta \equiv \det(\alpha_{rs}) \neq 0 \qquad (2.6)$$

A second set of conditions is obtained by the requirement that the solutions for μ_1 and μ_2 obtained from (2.4) must be positive. Thus we have the two diophantine equations

$$\frac{1}{\Delta}\begin{vmatrix} \lambda_1 & \alpha_{12} \\ \lambda_2 & \alpha_{22} \end{vmatrix} > 0$$

$$\frac{1}{\Delta}\begin{vmatrix} \alpha_{11} & \lambda_1 \\ \alpha_{21} & \lambda_2 \end{vmatrix} > 0 . \qquad (2.7)$$

The coefficients α_{rs} are subject to satisfy the property

$$\alpha_{11} + \alpha_{21} = 1 = \alpha_{12} + \alpha_{22} \qquad (2.8)$$

which follows from (2.5).

Appendix II

These equations have in general many solutions. It suffices to give an example:

Let $\lambda_1 = \frac{1}{3}$, $\lambda_2 = \frac{2}{3}$, $\alpha_{11} = \frac{3}{4}$, $\alpha_{12} = \frac{1}{4}$, $\alpha_{21} = \frac{1}{4}$, $\alpha_{22} = \frac{3}{4}$.

Then all the conditions are satisfied and there exist two projections Q_1 and Q_2 and two numbers $\mu_1 = \frac{1}{3}$ and $\mu_2 = \frac{2}{3}$, which satisfy (2.3).

1) P. Jordan, Z. f.Physik 80, 285 (1933); Gött.Nachr. 569 (1932); 209 (1933).

2) von Neumann, P. Jordan, and E. Wigner, Ann.Math. 35, 29 (1934).

3) I.E. Segal, Annals of Math. 48, 930 (1947).

4) I.E. Segal, Discussion remark at the conference of fundamental problems at Austin, Texas April 1970.

5) G. Birkhoff and J. von Neumann, Annals of Math. 37, 823 (1936).

6) G. Mackey, Mathematical Foundations of Quantum Mechanics, Benjamin, New York 1963.

7) C. Piron, Helv. Phys.Acta, 37, 439 (1964).

8) J.M. Jauch and C. Piron. Helv. Phys.Acta 42, 842 (1969).

9) R. Scalettar, Thesis, Cornell University (1959)

10) L.M. Loomis, Mem. Amer.Math.Soc. 18, 36 (1955)

11) G. Ludwig, Z. f. Physik 181, 233 (1964)

12) J.C.T. Pool, Thesis, State Univ. of Iowa (1963).

13) A. Einstein, B. Podolsky and N. Rosen, Phys.Rev. 47,777 (1935)

14) B. d'Espagnat, "Conceptions de la Physique Contemporaine" Hermann, Paris (1965).

15) ref[8)] Theorem and Corollary of Section 5.

16) ref.[8] Section 6.

17) J.-P. Eckmann and Ph.Ch. Zabey, Helv.Phys.Acta 42, 420 (1969)

18) cf. J.M. Jauch and F. Rohrlich, Theory of Photons and Electrons, esp. Sections 12-5 and 12-6 p.274 ff. Addison-Wesley 1956.

19) J.A. Wheller, Ann. N.Y. Acad.Sci. 48, 219 (1946)

20) R.C. Hanna, Nature 142, 332 (1948)

21) E. Bleuler, H.L. Bradt, Phys.Rev. 73, 1398 (1948)

22) C.S. Wu, I. Shaknow, Phys.Rev. 77, 136 (1950)

23) M.H.L. Pryce, J.C. Ward, Nature 160, 435 (1947)

24) H.S. Snyder, S. Posternack, J. Hornbostel, Phys.Rev.73 440 (1948)

25) D. Bohm and Y. Aharonov, Phys.Rev. 108, 1070 (1957)

26) J.M. Jauch and F. Rohrlich, Theory of Photons and Electrons, Addison Wesley (1956) esp. p.41.

27) P Jordan, Z. Physik, 44, 292 (1927)

28) J.S. Bell, Physics 1, 195 (1965)

29) H. Langhoff, Zeit.f.Physik 160, 186 (1960)

30) L. Kasday, Thesis, Columbia University 1970

31) J.S. Bell, Rev.Mod.Phys. 38, 453 (1966)

32) J.M. Jauch and C. Piron, Helv.Phys.Acta 36, 827 (1963)

33) J. von Neumann. The Mathematical Foundations of Quantum Mechanics; Princeton Univ. Press, N.Y. 1955 esp. p.295 ff.

34) D. Bohm and J. Bub, Rev.Mod. Phys. 38, 453 (1966)

35) J.M. Jauch, Foundations of Quantum Mechanics, Addison Wesley, 1968.

36) E.P. Wigner, Preprint, to be published

37) P.A.M. Dirac, Principles of Quantum Mechanics, Oxford 1955, p.7

38) F. London and E. Bauer, "La Théorie de l'Observation en Mécanique Quantique", Actualités Scientifiques et Industrielles, 775, Paris, Hermann (1939)

39) E. Wigner, The Scientist Speculates, p. 284, I.J. Good, ed. London: Heinemann (1962)

40) J.M. Jauch, E.P. Wigner and M.M. Yanase, Nuovo Cim. 48, 144 (1967).

41) M.D. Zeh, Found. of Physics, vol.1, no.1, p.67 (1970)

42) A. Daneri, A. Loinger and G.M. Prosperi, Nucl.Phys. 33, 297 (1962)

43) L. Rosenfeld, Progr. of Theor. Phys. Suppl. 1965

44) E.P. Wigner, Z. Physik 133, 101 (1952)

45) H. Araki and M.M. Yanase, Phys.Rev. 120, 622 (1960)

46) M.M. Yanase, Phys. Rev. 123, 666 (1961)

47) N. Bohr, Atomic Physics and Human Knowledge, Interscience, 1963, esp. p.3

48) J.M. Jauch, Helv.Phys. Acta 37, 293 (1964)