switching devices to separate positive from negative velocity data). The disadvantages are the requirement of a large-area absorber and the fact that the velocity is not uniform over the whole area.

The gamma-ray counting equipment is usually a scintillation or proportional counter in conjunction with a single-channel pulse-height analyzer and a scaler-timer.

See R. L. Mössbauer's original papers (item 4) for descriptions of a velocity spectrometer.

45. G. DePasquali, H. Frauenfelder, S. Margulies, and R. N. Peacock. Phys. Rev. Letters **4**, 71 (1960). Shows the inclined disk motion.

46. R. V. Pound and G. A. Rebka, Jr. Phys. Rev. Letters **3**, 554 (1959). Describes a motion based on an electromechanical transducer for which a loudspeaker may be substituted.

Further details on instrumentation may be found in items 11 and 12.

Some of the required equipment is available from the following suppliers:

Sources containing the radioisotope Co^{57} and absorbers containing stable Fe^{57}:
  Nuclear Science and Engineering
  P. O. Box 10901, Pittsburgh, Pennsylvania.

Thin-window scintillation crystals (x-ray type):
  Harshaw Chemical Company
  Cleveland, Ohio.

Low-noise photomultipliers (EMI 9536S):
  Hoffman Electron Tube Corporation
  Westbury, Long Island, New York.

Complete scintillation counter spectrometers are available from many suppliers, such as:
  Baird–Atomic, Inc., Cambridge, Massachusetts;
  Hamner Electronics Company, Inc., Princeton, New Jersey;
  Radiation Instrument Development Laboratory, Inc., Melrose Park, Illinois.

The Problem of Measurement

Eugene P. Wigner
Princeton University, Princeton, New Jersey

(Received 14 September 1962)

The standard theory of measurements in quantum mechanics is reviewed with special emphasis on the conceptual and epistemological implications. It is concluded that the standard theory remains the only one which is compatible with present quantum mechanics. Hence, if one wants to avoid the conclusion that quantum mechanics only gives probability connections between subsequent observations, the quantum-mechanical equations would have to be modified. Particular attention is paid to the case that the measuring apparatus is macroscopic and its state vector not accurately known before the measurement.

INTRODUCTION

THE last few years have seen a revival of interest in the conceptual foundations of quantum mechanics.[1] This revival was stimulated by the attempts to alter the probabilistic interpretation of quantum mechanics. However, even when these attempts turned out to be less fruitful than its protagonists had hoped,[2] the interest continued. Hence, after the subject had been dormant for more than two decades, we again hear discussions on the basic principles of quantum theory and the epistemologies that are compatible with it. As is often the case under similar circumstances, some of the early thinking had been forgotten; in fact, a small fraction of

[1] Some of the more recent papers on the subject are: Y. Aharonov and D. Bohm, Phys. Rev. **122**, 1649 (1961); Nuovo cimento **17**, 964 (1960); B. Bertotti, Nuovo cimento Suppl. **17**, 1 (1960); L. de Broglie, J. phys. radium **20**, 963 (1959); J. A. de Silva, Ann. Inst. Henri Poincaré **16**, 289 (1960); A. Datzeff, Compt. rend. **251**, 1462 (1960); J. phys. radium **21**, 201 (1960); **22**, 101 (1961); J. M. Jauch, Helv. Phys. Acta **33**, 711 (1960); A. Landé, Z. Physik **162**, 410 (1961); **164**, 558 (1961); Am. J. Phys. **29**, 503 (1961); H. Margenau and R. N. Hill, Progr. Theoret. Phys. **26**, 727 (1961); A. Peres and P. Singer, Nuovo cimento **15**, 907 (1960); H. Putnam, Phil. Sci. **28**, 234 (1961); M. Renninger, Z. Physik **158**, 417 (1960); L. Rosenfeld, Nature **190**, 384 (1961); F. Schlögl, Z. Physik **159**, 411 (1960); J. Schwinger, Proc. Natl. Acad. Sci. U. S. **46**, 570 (1960); J. Tharrats, Compt. rend. **250**, 3786 (1960); H. Wakita, Progr. Theoret. Phys. **23**, 32 (1960); **27**, 139 (1962); W. Weidlich, Z. Naturforsch **15a**, 651 (1960); J. P. Wesley, Phys. Rev. **122**, 1932 (1961). See also the articles of E. Teller, M. Born, A. Landé, F. Bopp, and G. Ludwig in *Werner Heisenberg und die Physik unserer Zeit* (Friedrich Vieweg und Sohn, Braunschweig, 1961).

[2] See the comments of V. Fock in the *Max Planck Festschrift* (Deutscher Verlag der Wissenschaften, Berlin, 1958), p. 177, particularly Sec. II.

it remains as yet unrediscovered in the modern literature. Equally naturally, some of the language has changed but, above all, new ideas and new attempts have been introduced. Having spoken to many friends on the subject which will be discussed here, it became clear to me that it is useful to review the standard view of the late "Twenties" and this will be the first task of this article. The standard view is an outgrowth of Heisenberg's paper in which the uncertainty relation was first formulated.[3] The far-reaching implications of the consequences of Heisenberg's ideas were first fully appreciated, I believe, by von Neumann,[4] but many others arrived independently at conclusions similar to his. There is a very nice little book, by London and Bauer,[5] which summarizes quite completely what I shall call the orthodox view.

The orthodox view is very specific in its epistemological implications. This makes it desirable to scrutinize the orthodox view very carefully and to look for loopholes which would make it possible to avoid the conclusions to which the orthodox view leads. A large group of physicists finds it difficult to accept these conclusions and, even though this does not apply to the present writer, he admits that the far-reaching nature of the epistemological conclusions makes one uneasy. The misgivings, which are surely shared by many others who adhere to the orthodox view, stem from a suspicion that one cannot arrive at valid epistemological conclusions without a careful analysis of the *process of the acquisition of knowledge*. What will be analyzed, instead, is only the type of information which we can acquire and possess concerning the external inanimate world, according to quantum-mechanical theory.

We are facing here the perennial question whether we physicists do not go beyond our competence when searching for philosophical truth. I believe that we probably do.[6] Nevertheless, the ultimate implications of quantum theory's formulation of the laws of physics appear interesting even if one admits that the conclusions to be arrived at may not be the ultimate truth.

THE ORTHODOX VIEW

The possible states of a system can be characterized, according to quantum-mechanical theory, by state vectors. These state vectors—and this is an almost verbatim quotation of von Neumann—change in two ways. As a result of the passage of time, continuously, according to Schrödinger's time-dependent equation—this equation will be called the equation of motion of quantum mechanics. The state vector changes, however, also discontinuously, according to probability laws, if a measurement is carried out on the system. This second type of change is often called the reduction of the wavefunction. It is this reduction of the state vector which is unacceptable to many of our colleagues.

The assumption of two types of changes of the state vector is a strange dualism. It is good to emphasize at this point that the dualism in question has little to do with the oft-discussed wave-versus-particle dualism. This latter dualism is only part of a more general pluralism or even "infinitesilism" which refers to the infinity of noncommuting measurable quantities. One can measure the position of the particles, or one can measure their velocity, or, in fact, an infinity of other observables. The dualism here discussed is a true dualism and refers to the *two* ways in which the state vector changes. It is also worth noting, though only parenthetically, that the probabilistic aspect of the theory is almost diametrically opposite to what ordinary experience would lead one to expect. The place where one expects probability laws to prevail is the change of the system with time. The interaction of the particles, their collisions, are the events which are ordinarily expected to be governed by statistical laws. This is not at all the case

[3] W. Heisenberg, Z. Physik **43**, 172 (1927), also his article in *Niels Bohr and the Development of Physics* (Pergamon Press, London, 1955); N. Bohr, Nature **121**, 580 (1928); Naturwissenschaften **17**, 483 (1929) and particularly *Atomic Physics and Human Knowledge* (John Wiley & Sons, Inc., New York, 1958).

[4] See J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Verlag Julius Springer, Berlin, 1932), English translation by the Princeton University Press, Princeton, New Jersey, 1955. See also P. Jordan, *Anschauliche Quantentheorie* (Julius Springer, Berlin, 1936), Chapter V.

[5] F. London and E. Bauer, *La Théorie de l'observation en mécanique quantique*" (Hermann et Cie., Paris, 1939); or E. Schrödinger, Naturwissenschaften **23**, 807 ff. (1935); Proc. Cambridge Phil. Soc. **31**, 555 (1935).

[6] This point is particularly well expressed by H. Margenau, in the first two sections of the article in Phil. Sci. **25**, 23 (1958).

here: the uncertainty in the behavior of a system does not increase in time if the system is left alone, that is, if it is not subjected to measurements. In this case, the properties of the system, as described by its state vector, change causally, no matter what the period of time is during which it is left alone. On the contrary, the phenomenon of chance enters when a measurement is carried out on the system, when we try to check whether its properties did change in the way in which our causal equations told us they would change. However, the extent to which the results of all possible measurements on the system can be predicted does not decrease, according to quantum-mechanical theory, with the time during which the system was left alone; it is as great right after an observation as it is a long time thereafter. The uncertainty of the result, so to say, increases with time for some measurements just as much as it decreases for others. The Liouville theorem is the analog for this in classical mechanics. It tells us that, if the point which represents the system in phase space is known to be in a finite volume element at one given time, an equally large volume element can be specified for a given later time which will then contain the point representing the state of the system. Similarly, the uncertainty in the result of the measurement of Q, at time 0, is exactly equal to the uncertainty of the measurement of $Q_t=\exp(-iHt/\hbar)Q_0\exp(iHt/\hbar)$ at time t. The information which is available at a later time may be less valuable than the information which was available on an earlier state of the system (this is the cause of the increase of the entropy); in principle, the amount of information does not change in time.

CONSISTENCY OF THE ORTHODOX VIEW

The simplest way that one may try to reduce the two kinds of changes of the state vector to a single kind is to describe the whole process of measurement as an event in time, governed by the quantum-mechanical equations of motion. One might think that, if such a description is possible, there is no need to assume a second kind of change of the state vector; if it is impossible, one might conclude, the postulate of the measurement is incompatible with the rest of quantum mechanics. Unfortunately, the situation will turn out not to be this simple.

If one wants to describe the process of measurement by the equations of quantum mechanics, one will have to analyze the interaction between object and measuring apparatus. Let us consider a measurement from the point of view of which the "sharp" states are $\sigma^{(1)}$, $\sigma^{(2)}\cdots$. For these states of the object the measurement will surely yield the values $\lambda_1, \lambda_2, \cdots$, respectively. Let us further denote the initial state of the apparatus by a, then, if the initial state of the system was $\sigma^{(\nu)}$, the total system—apparatus plus object—will be characterized, before they come into interaction, by $a\times\sigma^{(\nu)}$. The interaction should not change the state of the object in this case and hence will lead to

$$a\times\sigma^{(\nu)}\rightarrow a^{(\nu)}\times\sigma^{(\nu)}. \qquad (1)$$

The state of the object has not changed, but the state of the apparatus has and will depend on the original state of the object. The different states $a^{(\nu)}$ may correspond to states of the apparatus in which the pointer has different positions, which indicate the state of the object. The state $a^{(\nu)}$ of the apparatus will therefore be called also "pointer position ν." The state vectors $a^{(1)}$, $a^{(2)}$, . . . are orthogonal to each other—usually the corresponding states can be distinguished even macroscopically. Since we have considered, so far, only "sharp" states, for each of which the measurement in question surely yields one definite value, no statistical element has yet entered into our considerations.[7]

Let us now see what happens if the initial state of the object is not sharp, but an arbitrary linear combination $\alpha_1\sigma^{(1)}+\alpha_2\sigma^{(2)}+\cdots$. It then *follows* from the linear character of the quantum-mechanical equation of motion (as a result of the so-called superposition principle) that the state vector of object-plus-apparatus after the measurement becomes the right side of

$$a\times[\sum\alpha_\nu\sigma^{(\nu)}]\rightarrow\sum\alpha_\nu[a^{(\nu)}\times\sigma^{(\nu)}]. \qquad (2)$$

Naturally, there is no statistical element in this result, as there cannot be. However, in the state (2), obtained by the measurement, there is a

[7] The self-adjoint (Hermitean) character of every observable can be derived from Eq. (1) and the unitary nature of the transformation indicated by the arrow. Cf. E. Wigner, Z. Physik **133**, 101 (1952), footnote 2 on p. 102.

statistical correlation between the state of the object and that of the apparatus: the simultaneous measurement on the system—object-plus-apparatus—of the two quantities, one of which is the originally measured quantity of the object and the second the position of the pointer of the apparatus, always leads to concordant results. As a result, one of these measurements is unnecessary: The state of the object can be ascertained by an observation on the apparatus. This is a consequence of the special form of the state vector (2), of not containing any $a^{(\nu)}\times\sigma^{(\mu)}$ term with $\nu\neq\mu$.

It is well known that statistical correlations of the nature just described play a most important role in the structure of quantum mechanics. One of the earliest observations in this direction is Mott's explanation of the straight track left by the spherical wave of outgoing α particles.[8] In fact, the principal conceptual difference between quantum mechanics and the earlier Bohr-Kramers-Slater theory is that the former, by its use of configuration space rather than ordinary space for its waves, allows for such statistical correlations.

Returning to the problem of measurement, we see that our alternatives either of conflict between the theory of measurement and the equations of motion, or an explanation of that theory in terms of the equations of motion, have been cleverly dodged. The equations of motion permit the description of the process whereby the state of the object is mirrored by the state of an apparatus. The problem of a measurement on the object is thereby transformed into the problem of an observation on the apparatus. Clearly, further transfers can be made by introducing a second apparatus to ascertain the state of the first, and so on. However, the fundamental point remains unchanged and a full description of an observation must remain impossible since the quantum-mechanical equations of motion are causal and contain no statistical element, whereas the measurement does.

It should be admitted that when the quantum theorist discusses measurements, he makes many idealizations. He assumes, for instance, that the measuring apparatus will yield some result, no matter what the initial state of the object was. This is clearly unrealistic since the object may move away from the apparatus and never come into contact with it. More importantly, he has appropriated the word "measurement" and used it to characterize a special type of interaction by means of which information can be obtained on the state of a definite object. Thus, the measurement of a physical constant, such as cross section, does not fall into the category called "measurement" by the theorist. His measurements answer only questions relating to the ephemeral state of a physical system, such as, "What is the x component of the momentum of this atom?" On the other hand, since he is unable to follow the path of the information until it enters his, or the observer's, mind, he considers the measurement completed as soon as a statistical relation has been established between the quantity to be measured and the state of some idealized apparatus. He would do well to emphasize his rather specialized use of the word "measurement."

This will conclude the review of the orthodox theory of measurement. As was mentioned before, practically all the foregoing is contained, for instance, in the book of London and Bauer.[5]

[8] N. F. Mott, Proc. Roy. Soc. (London) **126**, 79 (1929).

CRITIQUES OF THE ORTHODOX THEORY

There are attempts to modify the orthodox theory of measurement by a complete departure from the picture epitomized by Eqs. (1) and (2). The only attempts of this nature which will be discussed here presuppose that the result of the measurement is not a state vector, such as (2), but a so-called mixture, namely *one* of the state vectors

$$a^{(\mu)}\times\sigma^{(\mu)}, \tag{3}$$

and that this particular state vector will emerge from the interaction between object and apparatus with the probability $|\alpha_\mu|^2$. If this were so, the state of the system would not be changed when one ascertains—in some unspecified way—which of the state vectors (3) corresponds to the actual state of the system, one would merely "ascertain which of various possibilities has occurred." In other words, the final observation only increases our knowledge of the system; it does not change anything. This is not true if the state vector, after the interaction between

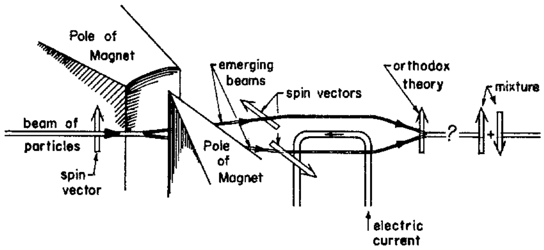

FIG. 1.

object and apparatus, is given by (2) because *the state represented by the vector (2) has properties which neither of the states (3) has.* It may be worthwhile to illustrate this point, which is fundamental though often disregarded, by an example.

The example is the Stern-Gerlach experiment[9] in which the projection of the spin of an incident beam of particles, into the direction which is perpendicular to the plane of the drawing, is measured. (See Fig. 1.) The index ν has two values in this case; they correspond to the two possible orientations of the spin. The "apparatus" is that positional coordinate of the particle which is also perpendicular to the plane of the drawing. If this coordinate becomes, in the experiment illustrated, positive, the spin is directed toward us; if it is negative, the spin is directed away from us. The experiment illustrates the statistical correlation between the state of the "apparatus" (the position coordinate) and the state of the object (the spin) which we have discussed. The ordinary use of the experiment is to obtain the spin direction, by observing the position, i.e., the location of the beam. The measurement is, therefore, as far as the establishment of a statistical correlation is concerned, complete when the particle reaches the place where the horizontal spin arrows are located.

What is important for us, however, is the right side of the drawing. This shows that the state of the system—object-plus-apparatus (spin and positional coordinates of the particle, i.e., the whole state of the particle)—shows characteristics which neither of the separated beams alone would have. If the two beams are brought together by the magnetic field due to the current in the cable indicated, the two beams will interfere and the spin will be vertical again. This could be verified by letting the united beam pass through a second magnet which is, however, not shown on the figure. If the state of the system corresponded to the beam toward us, its passage through the second magnet would show that it has equal probabilities to assume its initial and the opposite directions. The same is true of the second beam which was deflected away from us. Even though the experiment indicated would be difficult to perform, there is little doubt that the behavior of particles and of their spin conforms to the equations of motion of quantum mechanics under the conditions considered. Hence, the properties of the system, object plus apparatus, is surely correctly represented by an expression of the form (2) and shows, *in this case*, properties which are different from those of *either* alternative (3).

In the case of the Stern-Gerlach experiment, one can thus point to a specific and probably experimentally realizable way to distinguish between the state vector (2), furnished by the orthodox theory, and the more easily visualizable mixture of the states (3) which one would offhand expect. There is little doubt that in this case the orthodox theory is correct. It remains remarkable how difficult it is, even in this very simple case, to distinguish between the two, and this raises two questions. The first of these is whether there is, in more complicated cases, a principle which makes the distinction between the state vector (2), and the mixture of the states (3), impossible. As far as is known to the present writer, this question has not even been posed seriously heretofore, and it will be considered in the present discussion also only obliquely. The second question is whether there is a continuous transition between (2) and the mixture of states (3) so that in simpler cases (2) is the result of the interaction between object and measuring apparatus, but in more complicated and more realistic cases the actual state of object-plus-apparatus more nearly resembles a mixture of the states (3). Again, this question can be investigated within the framework of quantum mechanics, or one can postulate deviations from the quantum-mechanical equations

---

[9] The same experiment was discussed recently from another point of view by H. Wakita, Progr. Theoret. Phys. **27**, 139 (1962).

of motion, in particular from the superposition principle.

More complicated and more realistic means in the present context that the measuring apparatus, the state of which is to be correlated with the quantity to be measured, is of such a nature that it is easy to measure *its* state, i.e., correlate it with the state of another "apparatus." If this is done, the state of that second "apparatus" will be correlated also to the state of the object. The ease of establishing correlations between the state of the apparatus which came into direct contact with the object and another "apparatus" is usually greatest if the first one is of macroscopic nature, i.e., complicated from the quantum-mechanical point of view. The ease with which the secondary correlations can be established is a direct measure of how realistically one can say that the measurement has been completed. Clearly, if the state of the apparatus which carried out the primary measurement is just as difficult to ascertain as the state of the object, it is not very realistic to say that the establishment of a correlation between its and the object's state is a fully completed measurement. Nevertheless, it is so regarded by the orthodox theory. The question which we pose is, therefore, whether it is consistent with the principles of quantum mechanics to assume that at the end of a realistic measurement the state of object-plus-apparatus is not a wavefunction, as given by (2), but a mixture of the states (3). We shall see that the answer is negative. Hence, the modification of the orthodox theory of measurement mentioned at the beginning of this section is not consistent with the principles of quantum mechanics.

Let us now proceed with the calculation. Even though this point is not usually emphasized, it is clear that, in order to obtain a mixture of states as a result of the interaction, the initial state must have been a mixture already.[10] This follows from the general theorem that the characteristic values of the density matrix are constants of motion. The assumption that the initial state of the system, object-plus-apparatus, is a mixture, is indeed a very natural one because the state vector of the apparatus, which is under the conditions now considered usually a macroscopic object, is hardly ever known. Let us assume, therefore, that the initial state of the apparatus is a mixture of the states $A^{(1)}, A^{(2)}, \ldots,$ the probability of $A^{(\rho)}$ being p_ρ. The vectors $A^{(\rho)}$ can be assumed to be mutually orthogonal. The equations of motion will yield, for the state $A^{(\rho)}$ of the apparatus and the state $\sigma^{(\nu)}$ of the object, a final state

$$A^{(\rho)}\times\sigma^{(\nu)} \rightarrow A^{(\rho\nu)}\times\sigma^{(\nu)}. \qquad (4)$$

Every state $A^{(1\nu)}, A^{(2\nu)}, \ldots,$ will indicate the same state $\sigma^{(\nu)}$ of the object, the position of the pointer is ν for all of these. For different ν, however, the position of the pointer is also different. It follows that the $A^{(\rho\nu)}$, for different ν, are orthogonal, even if the ρ are also different. On the other hand, $A^{(\rho\nu)}$ and $A^{(\sigma\nu)}$, for $\rho\neq\sigma$, are also orthogonal because $A^{(\rho\nu)}\times\sigma^{(\nu)}$ and $A^{(\sigma\nu)}\times\sigma^{(\nu)}$ are obtained by a unitary transformation from two orthogonal states, $A^{(\rho)}\times\sigma^{(\nu)}$ and $A^{(\sigma)}\times\sigma^{(\nu)}$ and the scalar product of $A^{(\rho\nu)}\times\sigma^{(\nu)}$ with $A^{(\sigma\nu)}\times\sigma^{(\nu)}$ is $(A^{(\rho\nu)},A^{(\sigma\nu)})$. Hence, the $A^{(\rho\nu)}$ form an orthonormal (though probably not complete) system

$$(A^{(\rho\nu)},A^{(\sigma\mu)})=\delta_{\rho\sigma}\delta_{\nu\mu}. \qquad (5)$$

It again follows from the linear character of the equation of motion that, if the initial state of the object is the linear combination $\sum \alpha_\nu\sigma^{(\nu)}$, the state of object-plus-apparatus will be, after the measurement, a mixture of the states

$$A^{(\rho)}\times\sum \alpha_\nu\sigma^{(\nu)} \rightarrow \sum_\nu \alpha_\nu[A^{(\rho\nu)}\times\sigma^{(\nu)}]=\Phi^{(\rho)}, \qquad (6)$$

[10] This point is disregarded by several authors who have rediscovered von Neumann's description of the measurement, as given by (1) and (2). These authors assume that it follows from the macroscopic nature of the measuring apparatus that if several values of the "pointer position" have finite probabilities [as is the case if the state vector is (2)], the state is necessarily a *mixture* (rather than a linear combination) of the states (3)—that is, of states in each of which the pointer position is definite (sharp). The argument given is that classical mechanics applies to macroscopic objects, and states such as (2) have no counterpart in classical theory. This argument is contrary to present quantum-mechanical theory. It is true that the motion of a macroscopic body can be adequately described by the classical equations of motion if its state has a classical description. That this last premise is, according to present theory, not fulfilled, is clearly, though in an extreme fashion, demonstrated by Schrödinger's cat-paradox (cf. reference 5). Further, the discussion of the Stern-Gerlach experiment, given in the text, illustrates the fact that there are, in principle, observable differences between the state vector given by the right side of (2), and the *mixture* of the states (3), each of which has a definite position. Proposals to modify the quantum-mechanical equations of motion so as to permit a mixture of the states (3) to be the result of the measurement even though the initial state was a state vector, will be touched upon later.

with probabilities p_ρ. This same mixture should then be, according to the postulate in question, equivalent to a mixture of orthogonal states

$$\Psi^{(\mu k)} = \sum_\rho x_\rho^{(\mu k)} [A^{(\rho\mu)} \times \sigma^{(\mu)}]. \qquad (7)$$

These are the most general states for which the originally measured quantity has a definite value, namely λ_μ, and in which this state is coupled with some state (one of the states $\sum_\rho x_\rho^{(\mu k)} A^{(\rho\mu)}$) with a pointer position μ. Further, if the probability of the state $\Psi^{(\mu k)}$ is denoted by $P_{\mu k}$, we must have

$$\sum_k P_{\mu k} = |\alpha_\mu|^2. \qquad (7a)$$

The $x_\rho^{(\mu k)}$ will naturally depend on the α.

It turns out, however, that a mixture of the states $\Phi^{(\rho)}$ cannot be, at the same time, a mixture of the states $\Psi^{(\mu k)}$ (unless only one of the α is different from zero). A necessary condition for this would be that the $\Psi^{(\mu k)}$ are linear combinations of the $\Phi^{(\rho)}$ so that one should be able to find coefficients u so that

$$\sum_\rho x_\rho^{(\mu k)} [A^{(\rho\mu)} \times \sigma^{(\mu)}] = \Psi^{(\mu k)} = \sum_\rho u_\rho \Phi^{(\rho)}$$
$$= \sum_{\rho\nu} u_\rho \alpha_\nu [A^{(\rho\nu)} \times \sigma^{(\nu)}]. \qquad (8)$$

From the linear independence of the $A^{(\rho\nu)}$ it then follows that

$$u_\rho \alpha_\nu = \delta_{\nu\mu} x_\rho^{(\mu k)}, \qquad (8a)$$

which cannot be fulfilled if more than one α is finite. It follows that it is not compatible with the equations of motion of quantum mechanics to assume that the state of object-plus-apparatus is, after a measurement, a mixture of states each with one definite position of the pointer.

It must be concluded that *measurements which leave the system object-plus-apparatus in one of the states with a definite position of the pointer cannot be described by the linear laws of quantum mechanics*. Hence, if there are such measurements, quantum mechanics has only limited validity. This conclusion must have been familiar to many even though the detailed argument just given was not put forward before. Ludwig, in Germany, and the present writer have independently suggested that the equations of motion of quantum mechanics must be modified so as to permit measurements of the aforementioned type.[11] These suggestions will not be discussed in detail because they are suggestions and do not have convincing power at present. Even though either may well be valid, one must conclude that the only known theory of measurement which has a solid foundation is the orthodox one and that this implies the dualistic theory concerning the changes of the state vector. It implies, in particular, the so-called reduction of the state vector. However, to answer the question posed earlier: yes, there is a continuous transition between the state vector (2), furnished by orthodox theory, and the requisite mixture of the states (3), postulated by a more visualizable theory of measurement.[11]

WHAT IS THE STATE VECTOR?

The state vector concept plays such an important part in the formulation of quantum-mechanical theory that it is desirable to discuss its role and the ways to determine it. Since, according to quantum mechanics, all information is obtained in the form of the results of measurements, the standard way to obtain the state vector is also by carrying out measurements on the system.[12]

In order to answer the question proposed, we shall first obtain a formula for the probability that successive measurements carried out on a system will give certain specified results. This formula will be given both in the Schrödinger and in the Heisenberg picture. Let us assume that n successive measurements are carried out on the system, at times $t_1, t_2, \ldots, t_n$. The operators of the quantities which are measured are, in the Schrödinger picture, $Q_1, Q_2, \ldots, Q_n$. The characteristic vectors of these will all be denoted by ψ with suitable upper indices. Similarly, the characteristic values will be denoted by

[11] See G. Ludwig's article "Solved and Unsolved Problems in the Quantum Mechanics of Measurement" (reference 1) and the present author's article in *The Scientist Speculates*, edited by J. Good (William Heinemann, London, 1962), p. 284.

[12] There are, nevertheless, other procedures to bring a system into a definite state. These are based on the fact that a small system, if it interacts with a large system in a definite and well-known state, may assume itself a definite state with almost absolute certainty. Thus, a hydrogen atom, in some state of excitation, if placed into a large container with no radiation in it, will almost surely transfer all its energy to the radiation field of the container and go over into its normal state. This method of preparing a state has been particularly stressed by H. Margenau.

q so that

$$Q_j\psi_\kappa{}^{(j)} = q_\kappa{}^{(j)}\psi_\kappa{}^{(j)} \tag{9}$$

The Heisenberg operators which correspond to these quantities, if measured at the corresponding times, are

$$Q_j{}^H = e^{iHt_j}Q_j e^{-iHt_j} \tag{10}$$

and the characteristic vectors of these will be denoted by $\varphi_\kappa{}^{(j)}$, where

$$\varphi_\kappa{}^{(j)} = e^{iHt_j}\psi_\kappa{}^{(j)}; \quad Q_j{}^H\varphi_\kappa{}^{(j)} = q_\kappa{}^{(j)}\varphi_\kappa{}^{(j)}. \tag{10a}$$

If the state vector is originally Φ, the probability for the sequence $q_\alpha{}^{(1)}, q_\beta{}^{(2)}, \ldots, q_\mu{}^{(n)}$ of measurement results is the absolute square of

$$(e^{-iHt_1}\Phi,\psi_\alpha{}^{(1)})(e^{-iH(t_2-t_1)}\psi_\alpha{}^{(1)},\psi_\beta{}^{(2)})\cdots (e^{-iH(t_n-t_{n-1})}\psi_\lambda{}^{(n-1)},\psi_\mu{}^{(n)}). \tag{11}$$

The same expression in terms of the characteristic vectors of the Heisenberg operators is simpler

$$(\Phi,\varphi_\alpha{}^{(1)})(\varphi_\alpha{}^{(1)},\varphi_\beta{}^{(2)})\cdots(\varphi_\lambda{}^{(n-1)},\varphi_\mu{}^{(n)}). \tag{11a}$$

It should be noted that the probability is not determined by the n Heisenberg operators $Q_j{}^H$ and their characteristic vectors: the *time order* in which the measurements are carried out enters into the result essentially. Von Neumann already derived these expressions as well as their generalizations for the case in which the characteristic values $q_\alpha{}^{(1)}, q_\beta{}^{(2)}, \ldots$ have several characteristic vectors. In this case, it is more appropriate to introduce projection operators for every characteristic value $q^{(j)}$ of every Heisenberg operator $Q_j{}^H$. If the projection operator in question is denoted by $P_{j\kappa}$, the probability for the sequence $q_\alpha{}^{(1)}, q_\beta{}^{(2)}, \ldots, q_\mu{}^{(n)}$ of measurement-results is

$$(P_{n\mu}\cdots P_{2\beta}P_{1\alpha}\Phi,\ P_{n\mu}\cdots P_{2\beta}P_{1\alpha}\Phi). \tag{12}$$

The expressions (11) or (11a) can be obtained also by postulating that the state vector became $\psi_\kappa{}^{(j)}$ when the measurement of $Q^{(j)}$ gave the result $q_\kappa{}^{(j)}$. Indeed, the statement that the state vector is $\psi_\kappa{}^{(j)}$ is only a short expression for the fact that the last measurement on the system, of the quantity $Q^{(j)}$, just carried out, gave the result $q_\kappa{}^{(j)}$. In the case of simple characteristic values the state vector depends only on the result of the last measurement and the future behavior of the system is independent of the more distant past history thereof. This is not the case if the characteristic value $q^{(j)}$ is multiple.

The most simple expression for the Heisenberg state vector, when the j'th measurement gave the value $q_\kappa{}^{(j)}$, is, in this case

$$P_{j\kappa}\cdots P_{2\beta}P_{1\alpha}\Phi, \tag{12a}$$

properly normalized. If, after normalization, the expression (12a) is independent of the original state vector Φ, the number of measurements has sufficed to determine the state of the system completely and a pure state has been produced. If the vector (12a) still depends on the original state vector Φ, and if this was not known to begin with, the state of the system is a mixture, a mixture of all the states (12a), with all possible Φ. Evidently, the measurement of a single quantity Q the characteristic values of which are all nondegenerate, suffices to bring the system into a pure state though it is not in general foreseeable which pure state will result.

We recognize, from the preceding discussion, that the state vector is only a shorthand expression of that part of our information concerning the past of the system which is relevant for predicting (as far as possible) the future behavior thereof. The density matrix, incidentally, plays a similar role except that it does not predict the future behavior as completely as does the state vector. We also recognize that *the laws of quantum mechanics only furnish probability connections between results of subsequent observations carried out on a system.* It is true, of course, that the laws of classical mechanics can also be formulated in terms of such probability connections. However, they can be formulated also in terms of objective reality. The important point is that the laws of quantum mechanics can be expressed only in terms of probability connections.

PROBLEMS OF THE ORTHODOX VIEW

The incompatibility of a more visualizable interpretation of the laws of quantum mechanics with the equations of motion, in particular the superposition principle, may mean that the orthodox interpretation is here to stay; it may also mean that the superposition principle will have to be abandoned. This may be done in the sense indicated by Ludwig, in the sense proposed by me, or in some third, as yet unfathomed sense. The dilemma which we are facing in this regard makes it desirable to review any possible conceptual weaknesses of the orthodox interpretation and the present, last, section will be devoted to such a review.

The principal conceptual weakness of the orthodox view is, in my opinion, that it merely abstractly postulates interactions which have the effect of the arrows in (1) or (4). For some observables, in fact for the majority of them (such as xyp_z), nobody seriously believes that a measuring apparatus exists. It can even be shown that no observable which does not commute with the additive conserved quantities (such as linear or angular momentum or electric charge) can be measured precisely and in order to increase the accuracy of the measurement one has to use a very large measuring apparatus. The simplest form of the proof heretofor was given by Araki and Yanase.[13] On the other hand, most quantities which we believe to be able to measure, and surely all the very important quantities such as position, momentum, fail to commute with all the conserved quantities so that their measurement cannot be possible with a microscopic apparatus. This raises the suspicion that the macroscopic nature of the apparatus is necessary in principle and reminds us that our doubts concerning the validity of the superposition principle for the measurement process were connected with the macroscopic nature of the apparatus. The joint state vector (2), resulting from a measurement with a very large apparatus, surely *cannot be distinguished* as *simply from a mixture* as was the state vector obtained in the Stern-Gerlach experiment which we discussed.[14]

A second, though probably less serious, difficulty arises if one tries to calculate the probability that the interaction between object and apparatus be of such nature that there exist states $\sigma^{(\nu)}$ for which (1) is valid. We recall that an interaction leading to this equation was simply postulated as the type of interaction which leads to a measurement. When I talk about the probability of a certain interaction, I mean this in the sense specified by Rosenzweig or by Dyson who have considered ensembles of possible interactions and defined probabilities for definite interactions.[15] If one adopts their definition (or any similar definition) the probability becomes zero for the interaction to be such that there are states $\sigma^{(\nu)}$ satisfying (1). The proof for this is very similar to that[16] which shows that the probability is zero for finding reproducing systems—in fact, according to (1), each $\sigma^{(\nu)}$ is a reproducing system. The resolution of this difficulty is presumably that if the system with the state vector a—that is, the apparatus—is very large, (1) can be satisfied with a very small error. Again, the large size of the apparatus appears to be essential for the possibility of a measurement.

The simplest and least technical summary of the conclusions which we arrived at when discussing the orthodox interpretation of the quantum laws is that these laws merely provide probability connections between the results of several consecutive observations on a system. This is not at all unreasonable and, in fact, this is what one would naturally strive for once it is established that there remains some inescapable element of chance in our measurements. However, there is a certain weakness in the word "consecutive" as this is not a relativistic concept. Most observations are not local and one will assume, similarly, that they have an irreducible extension in time, that is duration. However, the "observables" of the present theory are instantaneous, and hence unrelativistic, quantities. The only exceptions from this are the local field operators and we know, from the discussion of Bohr and Rosenfeld, how many extreme abstractions have to be made in order to describe their measurement.[17] This is not a reassuring state of affairs.

The three problems just discussed—or at least

[13] H. Araki and M. Yanase, Phys. Rev. **120**, 666 (1961); cf. also E. P. Wigner, Z. Physik **131**, 101 (1952).

[14] This point was recognized already by D. Bohm. See Section 22.11 of his *Quantum Theory* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1951).

[15] C. E. Porter and N. Rosenzweig, Suomalaisen Tiedeakatemian Toimotuksia **VI**, No. 44 (1960); Phys. Rev. **120**, 1698 (1960); F. Dyson, J. Math. Phys. **3**, 140, 157, 166 (1962). See also E. P. Wigner, *Proceedings of the Fourth Canadian Mathematics Congress* (University of Toronto Press, Toronto, 1959), p. 174.

[16] Cf. the writer's article in *The Logic of Personal Knowledge* (Routledge and Kegan Paul, London, 1961), p. 231.

[17] N. Bohr and L. Rosenfeld, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **12**, No. 8 (1933); Phys. Rev. **78**, 194 (1950); E. Corinaldesi, Nuovo cimento **8**, 494 (1951); B. Ferretti, *ibid.* **12**, 558 (1954).

two of them—are real. It may be useful, therefore, to re-emphasize that they are problems of the formal mathematical theory of measurement, and of the description of measurements by macroscopic apparatus. They do not affect the conclusion that a "reduction of the wave packet" (however bad this terminology may be) takes place in some cases. Let us consider, for instance, the collision of a proton and a neutron and let us imagine that we view this phenomenon from the coordinate system in which the center of mass of the colliding pair is at rest. The state vector is then, if we disregard the unscattered beam, in very good approximation (since there is only S-scattering present)

$$\psi(r_p,r_n)=\frac{1}{r}e^{ikr}w(r), \tag{13}$$

where $r=|r_p-r_n|$ is the distance of the two particles and $w(r)$ some very slowly varying damping function which vanishes for $r<r_0-\frac{1}{2}c$ and $r>r_0+\frac{1}{2}c$, where r_0 is the mean distance of the two particles at the time in question and c the coherence length of the beam. If a measurement of the momentum of one of the particles is carried out—the possibility of this is never questioned—and gives the result $\mathbf{p}$, the state vector of the other particle suddenly becomes a (slightly damped) plane wave with the momentum $-\mathbf{p}$. This statement is synonymous with the statement that a measurement of the momentum of the second particle would give the result $-\mathbf{p}$—as follows from the conservation law for linear momentum. The same conclusion can be arrived at also by a formal calculation of the possible results of a joint measurement of the momenta of the two particles.

One can go even further[18]: instead of measuring the linear momentum of one particle, one can measure its angular momentum about a fixed axis. If this measurement yields the value $m\hbar$, the state vector of the other particle suddenly becomes a cylindrical wave for which the same component of the angular momentum is $-m\hbar$. This statement is again synonymous with the statement that a measurement of the said component of the angular momentum of the second particle certainly would give the value $-m\hbar$. This can be inferred again from the conservation law of the angular momentum (which is zero for the two particles together) or by means of a formal analysis. Hence, a "contraction of the wave packet" took place again.

It is also clear that it would be wrong, in the preceding example,[18] to say that even before any measurement, the state was a mixture of plane waves of the two particles, traveling in opposite directions. For no such pair of plane waves would one expect the angular momenta to show the correlation just described. This is natural since plane waves are not cylindrical waves or, since (13) is a state vector with properties different from those of any mixture. The statistical correlations which are clearly postulated by quantum mechanics (and which can be shown also experimentally, for instance in the Bothe-Geiger experiment) demand in certain cases a "reduction of the state vector." The only possible question which can yet be asked is whether such a reduction must be postulated also when a measurement with a macroscopic apparatus is carried out. The considerations around Eq. (8) show that even this is true if the validity of quantum mechanics is admitted for all systems.

[18] See, in this connection, the rather similar situation discussed by A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).