Chapter 3

The Quantum State Vector and Physical Reality[1]

Peter G. Bergmann

Department of Physics, Syracuse University, Syracuse, N. Y., U.S.A.

Ensembles of physical systems obeying quantum laws can be constructed in a variety of ways. The properties of ensembles constructed in the most common way may be summarized conveniently in terms of a single state vector in Hilbert space, or of a density matrix, and this circumstance might suggest that the quantum state represents indeed the essence of the state in which an individual physical system finds itself, just as this essence is provided in non-quantum physics by the coordinates of a point in phase space. In this paper it is shown that the properties of more general ensembles cannot be summarized in terms of state vectors or density matrices. In view of the fact that the assertions of quantum theory generally refer to statistical properties of ensembles, not to individual systems, it is contended that the identification of individual systems with specific quantum states is questionable.

This paper represents the further development of a line of exploration that was begun several years ago[2]. This earlier paper was concerned with the claim by von Neumann[3] that quantum measurements always lead to an increase of the dispersion, and hence of the entropy of ensembles of quantum systems. This monotonic increase, we showed, depends on the manner in which the ensemble is constructed. If the selection of systems to form the ensemble depends exclusively on a screening preceding in time the measurements to be performed, then the very process of measurement introduces an uncontrolled interaction of the systems with the apparatus of the experimenter, and the increase in dispersion is required by the theory, as it is intuitively acceptable. But we showed, by constructing explicitly alternative ensembles, that for one class the entropy decreases monotonically in time as the result of measurements, whereas in more general situations the change in entropy is not monotonic in either direction.

[1] This work has been supported in part by the Air Force Office of Scientific Research and by the Aerospace Research Laboratories, Office of Aerospace Research.

[2] Aharonov, Y., P. Bergmann, and J. Lebowitz: Phys. Rev. **134**, B1410 (1964).

[3] Neumann, John von: Mathematical foundations of quantum mechanics. Princeton: Princeton University Press 1955. (Translated from the German original by R. T. Beyer.)

Incidental to this principal thesis we pointed out that the assignment of a quantum state to a given system appears ambiguous. The purpose of the present paper is to elaborate this argument. In the conventional quantum theory of measurement an ensemble that has been obtained by the screening of systems of a given type in accordance with the outcome of a measurement of an observable A, in which the ensemble is to consist, say, of all those systems in which A is found to have the numerical value a, one assigns to the members of this ensemble the state $|a\rangle$. (It is assumed that a is a non-degenerate eigenvalue, and similar assumptions will be made in all that follows.) That in a subsequent measurement of the observable B the eigenvalue b_k will be observed will occur in a percentage of cases given by the absolute square of the product of the two vectors $|a\rangle$ and $|b_k\rangle$,

$$P(b_k) = \langle a | b_k \rangle \langle b_k | a \rangle . \tag{1}$$

This expression is bilinear in $|b_k\rangle$ and its dual $\langle b_k|$. The coefficients of this bilinear form, the idempotent matrix $|a\rangle\langle a|$, depend only on the selection procedure employed in constructing the ensemble on which measurements are to be performed. This is why the state $|a\rangle$ appears to be characteristic for the ensemble as such, and an intrinsic property of its member systems. The replacement of the matrix $|a\rangle\langle a|$ by a more general density matrix in the event of partial incoherence is well known and does not need to concern us now. Suffice it to point out that the density matrix, also, is contingent on a screening procedure preceding in time all subsequent manipulation of the physical systems composing the ensemble, a procedure which is merely less rigorous than confinement to one non-degenerate eigenvalue of an observable.

In what follows we shall be concerned with generalizations of the screening procedures. To avoid verbal ambiguities, we shall denote all probabilistic assertions of quantum theory by the term "assertion" if no time sense is to be implied. The term "prediction" will be reserved for assertions concerning the outcome of measurements in the future, and "retrodiction" for assertions concerning past observations. The standard formula (1) refers to predictions in the narrow sense. It can be extended to apply to a series of consecutive measurements of observables $B, C, \ldots, F$. The frequency with which the eigenvalues $b_k, c_m, \ldots, f$ will be obtained is given by the expression:

$$P(b_k, c_m, \ldots, f) = \langle f | \ldots \rangle \ldots \langle c_m | b_k \rangle \langle b_k | a | \rangle \langle a | b_k \rangle \langle b_k | c_m \rangle \ldots \langle \ldots | f \rangle . \tag{2}$$

The expression (2) is still part and parcel of the conventional approach to the quantum theory of measurement in that the screening observation of A precedes in time all other determinations, about which the probabilistic assertions are being made. It incorporates, among other facets, HEISENBERG's uncertainty relation, in that the probability of obtaining the specified sequence of results cannot be zero or one if adjacent observables have non-vanishing commutators. Incidentally, in order to avoid intermediate transition matrices, Eqs. (1) and (2) have been written down for observables that are constants of the motion, though not necessarily time-independent. As any observable can be converted trivially into a (time-dependent) constant of the motion, this restriction is formal rather than substantive.

Expression (2) may now be employed to obtain assertions referring to ensembles that have been obtained by means of multiple screening procedures. All the essentials are exhibited by an ensemble that consists exclusively of systems in which the initial measurement, of A, has led to the value a, and the final measurement, of F, to the value f. This ensemble is clearly a subensemble of the one screened exclusively with respect to the initial measurement, of A. Assertions concerning the intermediate observables $B, C, \ldots$ may be obtained readily by the standard techniques of contingent probabilities. For these frequencies we get the expression

$$\begin{aligned} P(b_k, c_m, \ldots) &= D^{-1} \langle f | \ldots \rangle \ldots \langle c_m | b_k \rangle \langle b_k | a \rangle \langle a | b_k \rangle \ldots \langle \ldots | f \rangle, \\ D &= \sum_{k'} \sum_{m'} \ldots \langle f | \ldots \rangle \ldots \langle c_{m'} | b_{k'} \rangle \langle b_{k'} | a \rangle \langle a | b_{k'} \rangle \ldots \langle \ldots | f \rangle. \end{aligned} \tag{3}$$

This expression is not multilinear in terms of the kets and bras corresponding to the possible outcomes of the intermediate observations, because of the denominator. Nor can it be written as the product of two factors one of which depends only on the screening procedures, the other only on the observations made on the resulting ensemble. Hence there is no simple manner in which some expression may be considered to be fully descriptive of the ensemble that has been constructed.

Because of the importance of the ensemble resulting from multiple-time screening for the argument of this paper, it should be emphasized that such ensembles are not without interest in experimental work. Particularly in high-energy and elementary particle research, samples for further analysis are frequently selected, by human scanning or by fully automated procedures (e. g. master pulse techniques), with respect to characteristics of both initial and final situations.

Ensembles constructed by two-time screening are not theoretically inferior to those obtained by the more conventional procedures of pre-selection. Ordinary quantum theory furnishes probabilistic formulas that permit one to make assertions concerning the outcome of experiments performed on these ensembles. But in some important respects

they have properties not shared by pre-selected ensembles. Consider, for instance, a two-time selected ensemble with two intermediate observations. If the observables corresponding to these intermediate observations do not commute with each other, it does not automatically follow that the product of their respective dispersions obeys HEISENBERG's uncertainty relation. Let, as an extreme example, the two intermediate observables, B and C, be identical with A and F, respectively, with respect to which the ensemble was screened,

$$B \equiv A, \quad C \equiv F, \quad [B, C] \neq 0. \tag{4}$$

The value of B will then be found to equal a, and that of C to be f, for all members of the ensemble. Hence the dispersion of the observations on B alone, and of those on C alone, will separately vanish, in apparent contradiction to the HEISENBERG relations. The real import of this result is, of course, not a logical contradiction but merely an injunction to be circumspect in the formulation of the uncertainty relations, which hold but for a limited class of ensembles.

I shall now return to the various roles that are played by the state vector of a quantum theoretical ensemble. First, it is part of the characterization of a type of physical system that one must construct the particular Hilbert space in which its observables are defined formally as Hermitian operators. This Hilbert space, being a linear vector space, defines a class of mathematical objects that are its constituent vectors, regardless of what additional physical meanings may be assigned to them. The present investigation changes nothing in this respect.

In the Schrödinger picture the kets are also the carriers of the dynamical law of the physical system; this role is not sacrosanct, considering the ease with which the dynamics can be fastened on the observables, in the Heisenberg picture, which not only resembles in this respect more closely the approach of classical mechanics but which, according to P. A. M. DIRAC[1], is preferable at least in some respects in quantum electrodynamics.

There remains the importance of the state of a system as a particularizing characteristic that sets it apart from other systems possessing the same structure (i. e. dynamical law) but happening to be "in a different state". This aspect of the state vector is shared by the Schrödinger and Heisenberg pictures as well as by other intermediate pictures. That the state vector has this function, the function carried in classical mechanics by the representative point in phase space (i. e. GIBBS' "phase"), is suggested by the circumstance that into the predictions concerning conventionally constructed ensembles the state vector of the systems enters as the only particularizing piece of information that is independent of the choice of the measurements to be performed. Surely, the state

[1] DIRAC, P. A. M.: Several public lectures, 1965 and 1966.

vector, or its generalization, the density matrix, are appropriate means for the characterization of ensembles that have been constructed by means of one-time screening prior to the onset of measurements.

For ensembles constructed by means of terminal one-time screening the state vector also is a proper means of characterization, and the only piece of information required for retrodiction. We have pointed out in Ref. 1 that in this case the ensemble is in the "state" established by subsequent screening, whereas in the conventional ensemble the state depends on the preceding screening. Two-time screening confronts us not with mere ambiguity; these ensembles simply cannot be described in terms of any one state vector or density matrix. This negative result tends in my opinion to undermine the significance of the state vector for the individual system as well.

The assignment of a state vector to a physical system operationally implies no more than the prediction of frequencies of outcomes of observations to be made subsequent to the establishment of a specified state. These predictions then refer not to individual physical systems but to ensembles composed of such systems, and moreover only to the class of ensembles constructed by means of one-time screening. Of course, ensembles are not objects occurring in nature but constructs of our intellect; what occurs in nature are individual physical systems. Systems may be imbedded in conceptual ensembles, or sets of real systems may be combined for purposes of discussion into ensembles. Experimental verification of the statistical assertions of quantum theory is based on the latter procedure. Because of the specialized character of one-time screened ensembles it appears open to question whether a means of characterizing such ensembles should be considered an intrinsic property of the component physical systems.

If the assignment of quantum states to individual physical systems is abandoned, the problematics of the change of quantum state as a result of the performance of measurements, i. e. the so-called collapse of the wave packet, evaporates. The performance of a measurement, and the use of its outcome as a means for additional screening, replaces the original ensemble by a subensemble, with its own statistical dispersion. The apparent resolution of the problems surrounding the extradynamic change in the quantum state of a system should, however, not be considered an unmitigated triumph. This advance is achieved only at the price of having to give up what hitherto has been considered the principal particularizing property of a quantum-theoretical physical system; no substitute is in sight.

Certainly, I do not consider that the discussion presented here leads to any definite conclusions. I believe that it calls attention to a possible novel point of view, whose import remains to be assessed.

Chapter 4

Probabilities in Quantum Mechanics[1]

Henry Margenau and **Leon Cohen**[2]

Department of Physics, Yale University, New Haven, Connecticut, U.S.A.

This note seeks to survey crucial issues and to pinpoint difficulties in the current controversy over the meaning of measurement [*1*] and its proper mathematical description in quantum physics. It deals primarily with those philosophic and mathematical considerations which arise from the statistical nature of the results of physical measurements. The first part reviews and aims to clarify matters of more or less general interest; the second and third are devoted to some specific recent results concerning joint probabilities and phase-space distribution functions.

I. One of the important features of quantum mechanical description is the scattering, the dispersion of measured values in repeated observations of a given quantity even when the physical state (ψ-function) of the system, which is the carrier of the quantity, is as precise and determinate as human ability permits. This fact has occasioned a variety of philosophical explanations: BOHR's and HEISENBERG's early belief that the statistical dispersion inherent in a quantum mechanical state is the reflection of the uncertainties introduced by the necessary interaction with the measuring apparatus which occurs at the time of measurement; DE BROGLIE's and BOHM's suggestion that obscure factors not appearing in the analysis (hidden variables) account for these fluctuations; HEISENBERG's later appeal to the Aristotelean theory of potentia [*2*], according to which the measurement of an observable in a state other than its eigenstate converts a potential quantity into an actually existing one; the distinction between "possessed" and "latent" observables [*3*], which formulates a new version of the philosophic contrast between primary and second ary qualities.

[1] Work supported by the United States Air Force Office of Scientific Research under Grant AFOSR 843/65.

[2] Now at Hunter College of the City University of New York.

The scattering of observations, and hence the use of probabilities, is of course not confined to quantum mechanics. Even in classical physics, analysis of the measurement process is meaningless without a theory of errors. The distinction between the two fields is thought to reside in the circumstance that the probabilities of classical physics are "reducible" [4], whereas those of quantum mechanics are not. That is to say, classical physics is in principle compatible with the assumption that knowledge of a system's state can be so refined that every conceivable physical quantity has an exact value, one which recurs in every measurement, so that the probability of the occurrence of any namable value is "reduced" either to zero or to one.

MAX BORN [5] has advanced an interesting argument which makes this statement a little artificial. Even in classical physics, he insists, the use of precise values of observables, unencumbered by probabilities, involves an idealization which defeats itself. It is surely not proper to assume exact knowledge of the instantaneous positions and velocities of all the molecules of a gas. A more reasonable premise is to assign to the values of these quantities small errors. This admission, however, opens the floodgates to ignorance, which develops in time. For the error attaching to position and velocity of a molecule increases in every collision. Roughly speaking, the relative error is doubled, and it is not difficult to see how long it will take for the error to grow to the magnitude of the quantity itself. If we start with an implausibly small error of one millionth of 1% in the knowledge of the original velocities, our knowledge will have been converted to ignorance in much less than one microsecond for a gas under standard conditions. In general, what physicists call relaxation time is also the time required for knowledge to be converted into ignorance in problems of this sort.

Such reasoning is of great practical importance, but it does not change the fact that classical physics, with its use of precise states, remains internally consistent and allows probabilities to be in principle reduced to certainties. Even the argument just cited is predicated upon the *functioning of exact laws* controlling the evolution of exact mechanical states whose incidental description or knowledge transcends our competence. No inherent principle of nature prevents us from ascertaining as exact a value as we please of the position and velocity of a single molecule at a given time, and there are ways in which we can study its behavior during an interval much smaller than the relaxation time, so that a dynamic account of its motion is possible. This leaves the distinction between quantum mechanics and classical physics as profound as ever; nor was it the intention of BORN, whose genius recognized it before all others, to deny it. His example serves to show forth the universal importance of probability.

Measurement is always a statistical affair. No single observation can be trusted to yield what goes as the "true" value of a physical quantity. Perhaps it is not amiss to sketch briefly how that "true" value is established.

Suppose we are making N measurements of a given quantity with the same apparatus. If N is sufficiently large, these results enable us to construct a curve, usually a Gauss error curve, which has a certain maximum and a certain half width. The half width is related to the probable error of the set of measurements, and the maximum of the curve defines the true value of the quantity. If the curve is indeed an error curve, then it may be shown that the arithmetical average of all the measured values are equal to the maximum of the curve. If N increases, the curve retains its width but it is possible to place the points more accurately upon it and thereby to define the maximum, the true value, more precisely. Hence, the probable error remains no matter how large the value of N.

To reduce the width of the curve different measuring apparatus must be employed and the more precise the measuring device the narrower the curve which results from the measurements. Thus arises the question: Is it possible within reasonable limits to choose different measuring instruments of increasing precision in such a way that the width of the curve becomes smaller and smaller but falls each time within the range of the preceeding grosser curve? For instance if we are to measure the length of a given object, the first device might be a carpenter's rule, the second a carefully calibrated yardstick, the third a Vernier caliper, the fourth a traveling microscope, the fifth an interferometer device, and so forth. In classical physics it had always been assumed that this kind of convergence, the nestling of successive error curves of smaller and smaller widths within the larger, was a fact of nature. Quantum mechanics revealed that there exists no set of measuring apparatus of progressive refinement which will make this kind of convergence true. In the fine-grain domain of atomic physics, errors cannot in general be confined to preassigned limits, and this is because of HEISENBERG's indeterminacy principle, which we shall now briefly consider.

In the form $\Delta p \cdot \Delta q \geqq \hbar/2$ the uncertainty principle is said to relate the "error" in the measured momentum p to the "error" in the position q of a particle. Physicists understand the application of this inequality very well and employ it without difficulty, but the explanation they give of it, especially in textbooks, is often highly questionable. Sometimes, major atrocities are committed in its name.

The most prevalent simple view concerning the origin of uncertainty in quantum mechanics holds the disturbance during the measurement process responsible for it. The γ-ray microscope carries much persuasive force in this line of reasoning: To measure q accurately one needs a short wave length γ-ray, and this produces, by way of the particle's recoil, an uncontrollable and unpredictable effect upon its momentum; hence a q-measurement entails an unavoidable error in p, and vice versa. Now all this, and the neat algebra that usually goes with the discussion, are certainly correct, but what they have to do with HEISENBERG's principle is far from clear. *Four very elementary difficulties emerge at once.*

The inequality asserts that in any state of the particle that is at any instant of time, and whether a measurement is being performed or not, the uncertainties or errors, cannot be smaller than it specifies. According to the disturbance theory, however, where Δp is due to the measurement interaction, it can only refer to the state of the particle *after* the measurement, while Δq relates to the position *before*. To contradict this inference would be to hold that measurements do not reveal what is, but what will be, or that they can play either role.

The second difficulty lies in the use of words like "unpredictable" and "uncontrollable". The disturbance theory has a decidedly classical cast; it pursues the interaction of the particle with the photon in complete classical detail, and if it were consistent it should admit both the controllability and the predicability of the results, for they are provided by electrodynamics. But here, as a concession which completely surrenders the point at issue, it interposes a hiatus and brands the interaction mysterious. The argument therefore begs the question.

Thirdly, the disturbance theory — unless it is taken as an inductive exemplification or merely as an illustration of the uncertainty principle — violates logic. Surely, quantum mechanics is a wider discipline than classical physics, based upon axioms additional to those which underlie the latter science. In other words, classical physics is a limiting case of quantum mechanics. Hence it can not possibly contain it logically since its range is smaller, and one cannot use classical reasoning to derive the uncertainty principle.

Finally, there is trouble with the definition of the Δ symbol. Errors, as we have seen, have no meaning with respect to a single measurement; being deviations from a mean they necessarily presuppose an ensemble of values. Thus, unless the γ ray microscope story is extended to embrace a great number of interactions between a particle and a photon and Δ is understood to refer to the statistical deviation of the results, the disturbance hypothesis is without relevance. If it is so extended, then it becomes unclear why only in the majority of instances, and not in all, an energetic photon causes a large disturbance, for there are results for which the deviation from the mean is very small. To us there appears no way in which all four difficulties can be removed.

Besides the disturbance theory, there is another popular simple conjecture which is invoked to explain uncertainty, again within the frame of the conventional pictorial concepts of mechanistic physics. Here the approach is operational; it is said, and made plausible by numerous examples, that it is instrumentally impossible to measure p and q simultaneously; the hardware clashes. This claim is expanded to cover all non-commuting observables: failure of operators to commute is supposed to symbolize experimental incompatibility of measuring procedures. When coupled with the doctrine of disturbance, the claim renders a rather pleasing account, for if p must invariably be measured before or after q, the effect of the interference can not be avoided.

The outlawing of simultaneous measurements, however, entrains its own peculiar infelicities. In the first place, there is no consideration whatever which prevents an experimenter from letting a particle reflect, at the same time or as nearly the same time as he pleases, a soft X-ray and a hard γ-ray. He can thus obtain numbers to be assigned to p and q, and this belies the thesis barring simultaneous measurements. To be sure, the simultaneous values for p and q thus obtained — whose errors escape our knowledge because there is no ensemble — can not be used as premises for causal prediction, they are without interest by themselves, and they do not define at quantum mechanical state. But they are available, and this makes the claim under examination palpably false.

But it exhibits a further, even more serious philosophic flaw. Science is the rational accomodation of contingent facts or observations. A theory which prescribes what can in the future be achieved or, worse, proscribes specific empirical acts, like simultaneous observations of p and q, transcends its own competence by restricting the development of experimental techniques. It becomes legislative when it should be explanatory. For these reasons it is necessary to look deeper into the bases of indeterminacy, and this forces us to renounce the encrusted habits of visual pictorization of elementary events, to relinquish the attempt to explain quantum uncertainty in terms of the familiar notions of conventional particle trajectories or wave propagation. Although it is heresy to say so, we believe that there is no dualism, no complementarity in quantum physics. The electron is neither a particle nor a wave, even if our accustomed language induces us to use these words. They are, strictly speaking, metaphorical and allude to something whose description exceeds the bounds of visual perception and is therefore obliged to discard such concepts as particles and waves, except as imperfect models of reality. Classical physics developed in the aegis of Cartesian clarity, its concepts answered the question: What can the mind's eye imagine? Quantum mechanics was born out of concerns characterized by the question: What can be measured? It has now left this area and has begum to ask: What can be conceived abstractly, but with mathematical and logical consistency?

After this digression we return to a correct, if minimal, interpretation of the uncertainty principle, which, though short of serving as an explanation, goes far in showing the inadequacy of the foregoing views. If one follows through the mathematical derivation of the formula, $\Delta p \cdot \Delta q \geqq \hbar/2$, the meaning of Δp and Δq becomes very clear; they represent the standard derivations, or the square root of the variances, of sets of repeated observations of p and q when the system (loosely speaking, the particle) is in an arbitrary but definite quantum state ψ. For example, we prepare the state of the system at time t, by passing it through a filter, or through a Stern-Gerlach field, or merely by waiting until an atoms has settled to its lowest energy state. Then, at time $t_1+\tau$, we perform a measurement of p, obtaining a value p_1. What happens to the system after $t_1+\tau$ depends on the manner in which the measurement was made, and knowledge of ψ does not in general permit its prediction. Different interactions may change ψ in many different ways; what is important from the point of view of the measurement is the number p_1.

To procedure an ensemble, the system must be reprepared at a later time t_2. A second measurement of p is then performed at $t_2+\tau$, and the result is p_2. In this way, by m repreparations and measurements, we obtain the set of numbers $p_1, p_2 \dots p_m$. By a corresponding procedure we can generate a set $q_1 \dots q_n$. The principle asserts that the standard deviations of these p's and these q's satisfy the inequality, for any ψ.

From this discussion it becomes apparent that the "error" in q, or Δq, can not be caused by the disturbance wrought in the measurement of any p, for the system was in each instance reprepared. Nor does it matter in what order the values appear, or in what manner the q-measurements are interspersed among those of p.

A succession of repeated preparations creates an ensemble, the instances of which are formed by one (perhaps the same, although this can rarely be guaranteed) system in the same state at different times. One might call this a time ensemble. Quantum mechanics is known to apply to space ensembles too. These are present when many systems, each in the same state, and isolated from each other, are present at the same time and the measuring device interacts with them all. A single observation will then yield the set $p_1 \dots p_m$, where m is now the number of systems present. Another single observation gives $q_1 \dots q_m$. Indeed these observations can also be performed at the same time. In practice, the space ensemble is of more common occurrence, but the uncertainty principle holds for both.

In the sequel we shall designate this just preceding account of the principle as its *standard version*. Whether it admits other interpretations is at the moment not clear.

We now examine briefly the logic of the probability concept as it is employed in quantum mechanics. As is well known, there are two extreme, basic approaches to the meaning of probability [*5*]. One, sometimes called a priori or subjective, establishes probability as a measure of the degree of confidence or expectation which a person may have in the outcome of an event. Its first mathematical formulation was given by LAPLACE, and one of its contemporary advocates is H. JEFFREY [*6*]. When viewed in this way, probability becomes in principle untestable and, at least in its most elementary sense, it changes with incidental evidence concerning the event. For instance the appearance of a number from 1 to 6 when a die is thrown is known to be $\frac{1}{6}$ before this event. If probability is lodged in one's knowledge with respect to the outcome of the next throw, its value changes from $\frac{1}{6}$ either to 1 or 0; the event "reduces" it to certainty. There are ways of escaping this conclusion even while maintaining the subjective interpretation, but they are not interesting here because it is just this conclusion, which springs from the most radical subjectivism in the logic of probability, that has left an imprint on the philosophy of quantum mechanics.

The principal representative of various objective views is the frequency theory, which holds probability to be the relative frequency of the occurrence of a specified event, or of a set of events, in a large aggregate of occurrences. The probability of throwing a 5 with a die is defined as the ratio of the number of times a 5 appears in a series of throws to the total number of throws. Although this ratio fluctuates as the number of throws increases, and indeed never attains a limit in the ordinary mathematical sense, it has become possible to define it with a precision sufficient for successful use.

It may not be out of place to emphasize here that probability in the latter sense, is a measurable quantity, entirely on a par with every other physical quantity. This needs to be recalled in order to counteract the superstition that probability stalks like a ghost through science, that its invocation is always a concession of ignorance. As a matter of fact, the assignment of probabilities to the numbers on the faces of a die, if they are understood in the objective manner, are attributes of the die which are as real as its color or its size or its weight. The reason people regard them otherwise is in the circumstance that probabilities can not be observed in one fell swoop, in a single measurement. However, as our earlier discussion shows, the other physical properties also require in essence a large number of observations in order that their "true" values be found. The psychological bias causing many to look upon probabilities as scientific citizens of second rank seems to stem from its lowly birth in minds preoccupied with games of chance.

One last point on the general logic of probability. It is wrong to view the subjective thesis as *opposed* to the objective one. Every physical theory has necessarily a purely formal aspect and an operational one. The former allows it to predict, the latter to confirm or refute its predictions [*5*]. The two logically disparate versions of probability here reviewed are therefore complementary to each other, and the very existence of these two required components is added evidence for the normal and respectable status of probability as a physical quantity.

Our discussion of the uncertainty principle, in its standard version, makes clear at once that the objective frequency interpretation is relevant for quantum mechanics. If the state function ψ has its postulated meaning, it too must be given a frequency sense. This has several consequences. First, since a relative frequency can never be determined by means of a single observation, since one throw of a die does not determine the next, there can be no presumption that a single observation of ψ, i. e. a measurement, *must establish its form*, or *modify it in a way fully known*, nor that is must determine the outcome of the next measurement. The other consequence is that a single observation does not change the probability in question at all. Objectively, whether or not a 2 appears in the next throw of a die, the probability of its appearance goes right on being $\frac{1}{6}$ afterwards.

Yet both of these conclusions are often violated when the meaning of ψ is discussed by physicists. VON NEUMANN introduced the so-called projection postulate; textbooks speak of the reduction of a wave packet upon measurement; these amount to the posit that a single measurement has the effect of producing an eigenstate, namely one corresponding to the measured value. This hypothesis has been shown to be untenable elsewhere (references 4, 5 and literature cited there) and its short comings need not be exhibited again. It might be of some interest, however, to speculate upon the reasons why this strange hypothesis has had such persuasive force.

One of them is surely psychological; the projection idea is a hangover from classical physics, where every measurement (in a very naive sense which is actually contradictory to the theory of errors) insures the same outcome upon repetition. There, every measurement is also the preparation of a state. In quantum mechanics this need not be the case.

The other reason is mathematical. Quantum mechanics associates operators with physical acts yielding numbers. Measurement itself represents the most universal act of this type, and it is tempting to seek an operator, M, which symbolizes it in universal fashion. Specific operators for specific measurements are already known. Now, what are the

properties of measurement in general? To say that it always generates a real number is not helpful in this search, for it yields an indiscriminate, large class of operators. One may look upon it in another way, however. A measurement puts a question to nature: does this variable have a value within this given range? The answer is either yes or no, it can be symbolized as 1 or 0. Hence the operator, M, must have two eigenvalues, M', namely 1 and 0, and no more. These satisfy the equation $M'^2 = M'$; hence M must satisfy it also. But the operator, defined by $M^2 = M$ is known to be the projection operator, constructible for any state ψ in well known ways.

This somewhat formal success of the search for M entails a mathematical conclusion. When M is applied to a vector in Hilbert space, say $\boldsymbol{x}$, it changes the direction of $\boldsymbol{x}$ and reduces its magnitude so that it becomes the component of $\boldsymbol{x}$ in a new direction. In short, it projects $\boldsymbol{x}$ upon some ray in Hilbert space. Physically, this means it creates a new state. Hence the projection hypothesis.

The cogency of such reasoning is of course spurious. Even if M is accepted as a valid symbol for a universal measurement, the application of M to an *arbitrary* $\boldsymbol{x}$ need not be interpreted as the outcome of measurement upon a system in state $\boldsymbol{x}$ at all. There is no warrant for this hypothesis in the axioms of quantum mechanics. More serious, however, is the identification of M with measurement in general. Why should there be a valid symbol of a specific sort for every possible measurement process? In classical physics we have formulas for momentum, kinetic and potential energy and all other important observables. But do we have, or *need*, a formula for an unspecified observable? Our suspicion is that it would be trivial if written down.

We therefore end this section with the assertion that quantum mechanics deals with measurable probabilities which take the form of relative frequencies; that the interaction of a physical system with a measuring device does not necessarily project its state onto an eigenvector in Hilbert space-but rather projects thath state into irrelevance. As a rule, quantum mechanics ceases to have interest in the state of a system after an observable has been measured upon it. There is no single axiom which defines the post-measurement condition. Nevertheless, there are operations (e. g. correlation experiments) which insure that a given observed value will always be associated, simultaneously or in sequel, with a specifiable other value, observed or not. These, however, are not governed by one basic axiom and require special treatment. They are best viewed as single but compound measurements [*7*].

II. In section I attention has been confined to the probability that, the state of a system being ψ, a measurement of an observable r shall give the eigenvalue r_i. It is specified by the postulates of quantum mechanics as

$$P(r'; \psi) = |\langle\psi| r_i\rangle|^2 \tag{1}$$

provided $\langle\psi| r_i\rangle$ is the scalar product of the bra vector $\langle\psi|$ and the ket vector $|r_i\rangle$, which is an eigenstate of r corresponding to r_i. In the preceding section we reviewed and criticised various hypotheses relating to the measurement of two different observables, say x and y (e. g. position and momentum), simultaneously or in sequence, and therefore the question arises as to the probability that, when a quantum system is in state ψ, a measurement of x and one of y shall yield the eigenvalues x_i and y_j. The cases in which $y \equiv x$, or $y \equiv (x$ at a later time) are included here. This introduces the concept of a joint probability, to be designated by $P(x_i, y_j; \psi)$. The ordinary postulates are non-committal with respect to it construction, and here lies the mathematical root for the disparity of views regarding what happens when p and q, for instance, are being measured. Some philosophers of physics argue in fact that a construct like $P(x_i, y_j; \psi)$ is meaningless except when the operators for x and y commute, usually by an appeal to the belief that joint measurements for them are impossible. They are then inclined to devise non-Aristotelean quantum logics which open unlimited vistas for speculation. We have already given reasons to doubt this conjecture. If it were true in the literal sense of every proper theory of stochastic variables, then it *ought* to be possible to construct a joint probability, and furthermore $P(x_i, y_j; \psi)$ should be zero for every i and j when x and y do not commute. At any rate, the search for possible formulations of a joint probability function must not be foregone. If none exists, that is at least worth knowing.

The search begins with a statement of the mathematical conditions which a joint probability must satisfy. These are, positiveness of P, (2) and (3, 4) the necessary relation to the marginal probabilities. They have the form

$$P(x_i, y_j; \psi) \geqq 0, \tag{2}$$

$$\sum_i P(x_i, y_j; \psi) = P(y_j, \psi) = |\langle\psi| y_j\rangle|^2, \tag{3}$$

$$\sum_j P(x_i, y_j; \psi) = P(x_i, \psi) = |\langle\psi| x_i\rangle|^2, \tag{4}$$

One may wish to impose further conditions, demanding that, when used in the calculation of expectation values in the customary way, they shall lead to the correct quantum mean; e. g. that

$$\mathrm{Exp}(x) = \sum_{ij} P(x_i, y_j; \psi)\, x_i = \langle\psi| \boldsymbol{x} |\psi\rangle,$$

where x_{op} is the operator corresponding to the observable x. This greatly encumbers the search. We return to this problem in the following section and restrict ourselves here to conditions 2, 3 and 4 which seem to suffice as a basis for a minimal theory of measurement.

A theorem of statistics relates the covariance of two random variables x and y to their joint probabilities in the following way.

$$\mathrm{Cov}(x y) = \sum_{ij} P(x_i, y_j)\, x_i\, y_j - \langle x\rangle\langle y\rangle. \tag{5}$$

As to notation, the covariance

$$\mathrm{Cov}(x y) \equiv \langle x y\rangle - \langle x\rangle\langle y\rangle$$

provided we use brackets to denote expectation values. Now quantum mechanics immediately suggests that we write for $\langle x y\rangle$ something like $\langle\psi| \boldsymbol{x}\boldsymbol{y} |\psi\rangle$, $\boldsymbol{x}$ and $\boldsymbol{y}$ being the operators corresponding to x and y, and if this is done one can extract the form of $P(x_i, y_j, \psi)$ from (5). Certainly $\langle x\rangle$ and $\langle y\rangle$ are known to be $\langle\psi| \boldsymbol{x} |\psi\rangle$ and $\langle\psi| \boldsymbol{y} |\psi\rangle$.

But the difficulty encountered here arises from the fact that the product xy can be translated into quantum mechanics in many different ways. What is needed is a *rule of correspondence* between a product like xy, or in general $x^n p^m$, in which the factors commute and which is directly observable, and its quantum equivalent. Here appears in very specific form the fundamental epistemological problem of the relation between direct experiences (P-plane of reference 5) and the constructs that symbolize them in our reasoning about the world (C-field). In classical physics that relation was held to be a trivial isomorphism — the construct temperature *was* the class of numbers provided by a thermometer; however here the rules of correspondence become a matter for deliberate choice and discrimination. More will be said about them in part III.

Suffice it here to record that a fairly obvious choice for $\boldsymbol{x}\boldsymbol{y}$ is obtained by symmetrization [8]:

$$\boldsymbol{x}\boldsymbol{y} \to \frac{\boldsymbol{x}\boldsymbol{y} + \boldsymbol{y}\boldsymbol{x}}{2}.$$

This leads, via Eq. (5) to the result for the joint probability

$$P(x_i, y_j; \psi) = R[\langle\psi|x_i\rangle\langle x_i| y_j\rangle\langle y_j|\psi\rangle] \tag{6}$$

where R stands for "real part". Formula (6) has certain attractive features. It provides a correlation coefficient $\sigma(x, y)$ defined by

$$\sigma(x, y) = \mathrm{Cov}(x y)\,[\mathrm{Var}(x)\ \mathrm{Var}(y)]^{-\frac{1}{2}} \tag{7}$$

which satisfies the desired inequality

$$-1 \leqq \sigma(x y) \leqq +1$$

for every x and y. It also conforms to Eqs. (3) and (4), but unfortunately it does not satisfy (2). P constructed in accordance with (6) is not in general positive definite, as closer inspection (and various examples worked out in reference 8) will show.

This one example, to which others will be added in the final section of this paper, places in view the major difficulty which afflicts nearly all formulations of $P(x_i, y_j; \psi)$. There may be a deeper reason for that, possibly hidden in strange and obscure features of the measurement process which make it similar to creation and annihilation in field theories[1]. At the moment, however, we shall take the stand that joint probabilities which violate (2) are to be rejected.

One then observes that there is *one* $P(x_i, y_j; \psi)$, a trivial one, which does obey all sum rules, (2—4), and which can be written down even without the detour over covariances. It is

$$P(x_i, y_j; \psi) = |\langle\psi|x_i\rangle|^2 \, |\langle\psi|y_j\rangle|^2. \tag{8}$$

While mathematically adequate from the present point of view, it has the fault of yielding no correlations between the random variables x and y: $\mathrm{Cov}(x y)$, constructed in conformity with (5) is zero, and so is $\sigma(x, y)$. Must one therefore reject it?

This may well be the case, but for reasons that are not yet wholly clear. It is noteworthy, however, that formula (5) contains exactly the proposition which was affirmed in our discussion of the standard version of uncertainty. There we saw that the probability of measuring p, *after repreparation of the state* on the time ensemble *or in simultaneous observations* on *a space ensemble*, was indeed independent of the measurement of q. Formula (8) describes that situation. It seems, therefore, that we are in possession of a reasonable theory of measurement, which endows joint probabilities with meaning, provided we accept the conditions noted in part I; they are: an interpretation of uncertainty as standard deviation of a set of measured values after repeated preparation of the state ψ, and the disavowal of a priori knowledge about the fate of the quantum system after measurement.

We now turn to the consideration of a special class of joint distribution functions, namely $P(q, p; \psi)$, where q and p continue to signify position and momentum. These are phase space distribution functions. Through ψ, the functional P depends also on the time t. In the next section we employ a notation which lends itself a little more naturally to the technical matters under discussion; we write

$$P(q, p; \psi(t)) = F(q, p, t)$$

[1] This point was emphasized to us in conversation by Professor VIGIER.

omitting the argument t when the time dependence is otherwise clear. Only one-dimensional systems will be treated; generalization to several dimensions is easy but entails more cumbersome equations. Finally, use of the ordinary notation, $\psi(q, t)$ in place of $\langle\psi|q, t\rangle$, is simpler and will therefore be adopted.

III. The possibility of formulating quantum mechanics in terms of ensembles in the phase space of position and momentum originated with WIGNER [*9*]. He found a function of position and momentum which satisfies Eqs. (3) and (4) for the case of q and p. The Wigner distribution is

$$F_w(q, p) = \frac{1}{2\pi} \int \psi^*\left(q - \frac{1}{2}\tau\hbar\right) e^{-i\tau p} \psi\left(q + \frac{1}{2}\tau\hbar\right) d\tau. \tag{9}$$

It can readily be verified that

$$\int F(q, p(\, dp = |\psi(q)|^2, \tag{10}$$

$$\int F(q, p) \, dq = |\varphi(p)|^2 \tag{11}$$

where $\varphi(p)$ is the momentum state function given by

$$\varphi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int \psi(q) \, e^{-i/\hbar \, q p} \, dq \tag{12}$$

and $|\varphi(p)|^2$ the quantum mechanical probability distribution for momentum. Although F_w does yield the correct marginal distributions, one cannot say that F_w is a true joint distribution of q and p because it does not in general satisfy (2). None the less, F_w can be formally used, to some extent, to calculate expectation values of quantum observables using phase space integration. A different distribution function (studied by MARGENAU and HILL [*8*]) is

$$F_1(q, p) = \frac{1}{4\pi} R\left\{\psi(q) \int e^{-i\tau q} \psi^*(q - \tau\hbar) \, d\tau\right\} \tag{13}$$

where R signifies the real part of the quantity in brackets. This represents another way of writing Eq. (6). If one assumes that position and momentum are not correlated, then we are lead to take distribution (8) which now reads

$$F_0(q, p) = |\psi(q)|^2 \, |\varphi(p)|^2. \tag{14}$$

Both (13) and (14) satisfy (10) and (11).

An explicit expression for the totality of functions of q and p which satisfy (10) and (11) can be given [*10, 11*]. It is

$$F(q, p, t; f) = \frac{1}{4\pi^2} \iiint e^{-i\vartheta q - i\tau p + i\vartheta u} \times f(\vartheta, \tau, t) \, \psi^*\left(u - \frac{1}{2}\tau\hbar\right) \psi\left(u + \frac{1}{2}\tau\hbar\right) d\vartheta \, d\tau \, du \tag{15}$$

where $f(\vartheta, \tau, t)$ is any function which satisfies

$$f(0, \tau, t) = f(\vartheta, 0, t) = 1. \tag{16}$$

Thus f is the key to an indefinite number of phase space distribution functions which can be generated via Eq. (15). That (15) is the set of all functions constrained by conditions (10) and (11) can be proved by first showing that a judicious choice of f will allow any function of q and p to be written in form (15)[1]. Then the integration of F with respect to p and q forces the imposition of (16) on f if the result of the integration is to yield (10) and (11). F_w, F_1 and F_0 are obtained from (15) by particularizing f to the value

$$f_w = 1 \tag{17}$$

$$f_1 = \cos \tfrac{1}{2} \vartheta \tau \hbar \tag{18}$$

$$f_0 = \frac{\iint |\psi(q)|^2 \, |\varphi(p)|^2 \, e^{i\vartheta q + i\tau p} \, dq \, dp}{\int \psi^*(u - \frac{1}{2}\tau\hbar) \, e^{i\vartheta u} \, \psi(u + \frac{1}{2}\tau\hbar) \, du} \tag{19}$$

respectively.

The transition from quantum to classical mechanics can be made very clear through the investigation of the time derivative of F, i. e. the "equation of motion" which F satisfies. If (15) is differentiated with respect to time and SCHRÖDINGER's equation is applied to ψ, one obtains, after a somewhat lengthy but direct calculation, the equation of motion for F.

$$\begin{aligned}\frac{\partial F(q, p, t; f)}{\partial t} = & \frac{\dot{f}\left(-i\frac{\partial}{\partial q_F}, -i\frac{\partial}{\partial p_F}\right)}{f\left(-i\frac{\partial}{\partial q_F}, -i\frac{\partial}{\partial p_F}\right)} F(q, p, t; f) \\ & + \frac{2}{\hbar} f^{-1}\left(i\frac{\partial}{\partial q_F}, i\frac{\partial}{\partial p_F}\right) f\left(-i\frac{\partial}{\partial q_H}, -i\frac{\partial}{\partial p_H}\right) \times \\ & \times \sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_F}\frac{\partial}{\partial q_H} - \frac{\partial}{\partial p_H}\frac{\partial}{\partial q_F}\right] \times \\ & \times f\left(i\frac{\partial}{\partial q_F} + i\frac{\partial}{\partial q_H}, i\frac{\partial}{\partial p_F} + i\frac{\partial}{\partial p_H}\right) H(q, p) \, F(q, p, t; f)\end{aligned} \tag{20}$$

where $\frac{\partial}{\partial q_H}, \frac{\partial}{\partial p_H}$ operate on H only and $\frac{\partial}{\partial q_F}, \frac{\partial}{\partial p_F}$ on F. $H(q, p)$ is the *classical* Hamiltonian. Now, suppose we choose f as a function of $\hbar$

[1] By taking the Fourier transform of (15) we have

$$f(\vartheta, \tau, t) = \frac{\iint F(q, p, t) \, e^{i\vartheta q + i\tau p} \, dq \, dp}{\int \psi^*(u - \frac{1}{2}\tau\hbar, t) \, e^{i\vartheta u} \, \psi(u + \frac{1}{2}\tau\hbar) \, du}.$$

Thus for any well-behaved F, we can find an f so that (15) is satisfied. We must, of course, also know φ.

such that

$$\lim_{\hbar \to 0} f \to 1, \tag{21}$$

$$\lim_{\hbar \to 0} \dot{f} \to 0. \tag{22}$$

Since

$$\sin\frac{\hbar}{2}(\;) \to \frac{\hbar}{2}(\;),$$

Eq. (20) becomes, in the limit as $\hbar$ approaches zero,

$$\frac{\partial F}{\partial t} = \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} - \frac{\partial H}{\partial p}\frac{\partial F}{\partial q} \tag{23}$$

which is LIOUVILLE's equation. The Wigner distribution and the distribution of MARGENAU and HILL satisfy (21) and (22). But the distribution which shows no correlations, that is defined in Eq. (14) which resulst from

$$f(\vartheta, \tau) = \frac{\iint |\psi(q)|^2 \, |\varphi(p)|^2 \, e^{i\vartheta q + i\tau p} \, dq \, dp}{\int \psi^*(u - \frac{1}{2}\tau\hbar) \, e^{i\vartheta u} \, \psi(u + \frac{1}{2}\tau\hbar) \, du} \tag{24}$$

does not in general satisfy (21) and (22). Hence F_0 does not obey the Liouville equation in the limit of $\hbar \to 0$. The reason for this failure is the fact that the Liouville equation reflects the presence of correlations while (14) admits no correlations in any limit.

As we have seen, the formation of the quantum mechanical operator from its classical counterpart is straightforward as long as the classical quantity is either a function of q or of p only, or if it is the sum of such functions. One merely replaces p by the usual operator $\boldsymbol{p}$ and q by $\boldsymbol{q}$. But if the classical observable contains product terms of q and p, then difficulties arise, for there is no unique way of forming the quantum mechanical operator. Several methods have been proposed for dealing with such cases. They are called rules of correspondence, and they were encountered and commented on in section II. The following rules are known and have been used.

a) DIRAC's rule:

$$\{\mathscr{A}, \mathscr{B}\} \to -\frac{i}{\hbar}[\boldsymbol{A}, \boldsymbol{B}] \tag{25}$$

where $\{\,,\,\}$ is the classical Poisson bracket of $\mathscr{A}$ and $\mathscr{B}$ and $[\,,\,]$ is the commutator of the operators $\boldsymbol{A}$ and $\boldsymbol{B}$

b) VON NEUMANN's rule:

If

$$\mathscr{A} \to \boldsymbol{A}$$

than for any function g

$$g(\mathscr{A}) \to g(\boldsymbol{A})$$

and if

$$\mathscr{A} \to \boldsymbol{A}$$

$$\mathscr{B} \to \boldsymbol{B}$$

then

$$\mathscr{A} + \mathscr{B} \to \boldsymbol{A} + \boldsymbol{B}$$

c) WEYL's rule:

$$q^n p^m \to \frac{1}{2^n} \sum_{l=0}^{n} \binom{n}{l} \boldsymbol{q}^{n-l} \boldsymbol{p}^m \boldsymbol{q}^l.$$

d) Rule of symmetrization:

$$q^n p^m \to \frac{1}{2} (\boldsymbol{q}^n \boldsymbol{p}^m + \boldsymbol{p}^m \boldsymbol{q}^n)$$

e) Rule of BORN and JORDAN:

$$q^n p^m \to \frac{1}{m+1} \sum_{l=0}^{m} \boldsymbol{p}^{m-l} \boldsymbol{q}^n \boldsymbol{p}^l.$$

The first two rules have been shown to be inconsistent [*12*].

There is an intimate connection between correspondence rules and the distribution functions F. As noted, this is due to the fact that a correspondence rule enables the calculation of the moments $\langle q^n p^m \rangle$, which in turn are generally sufficient to calculate the distribution function[1]. Hence we expect that for each distribution function given by (15), we can obtain a rule of correspondence. This is indeed the case. If $g(q, p)$ is the classical function, then it can be shown that the quantum mechanical operator, $G(\boldsymbol{q}, \boldsymbol{p})$ is given by

$$G(\boldsymbol{q}, \boldsymbol{p}) = \iint \gamma(\vartheta, \tau)\, f(\vartheta, \tau)\, e^{i\vartheta \boldsymbol{q} + i\tau \boldsymbol{p}} d\vartheta\, d\tau \tag{26}$$

where $\gamma(\vartheta, \tau)$ is the Fourier transform of $g(q, p)$ and $f(\vartheta, \tau)$ satisfies (16). An additional condition which must be imposed on f to assure that G is hermitian is

$$f(\vartheta, \tau) = f^*(-\vartheta, -\tau). \tag{27}$$

Proof of these statements can be found in references 10 and 11.

[1] This is done by forming the so-called characteristic function $M(\vartheta, \tau)$ from the moments,

$$M(\vartheta, \tau) = \sum_{n, m=0}^{\infty} \frac{(i\vartheta)^n (i\tau)^m}{n!\, m!} \langle q^n p^m \rangle.$$

The distribution function, F, is then

$$\frac{1}{4\pi^2} \iint M(\vartheta, \tau)\, e^{-i\vartheta q - i\tau p} d\vartheta\, d\tau.$$

See any standard text in probability theory.

Different choices of f yield different correspondence rules. Actually (26) generates all possible correspondence rules in the sense that once we have decided that

$$q^n \to \boldsymbol{q}^n \tag{28}$$

and

$$p^n \to \boldsymbol{p}^n \tag{29}$$

for all n, Eq. (26) gives the only possible choices of G which reduce to (28) or (29) when $g(q, p) = p^n$ or q^n. The Weyl rule, or the symmetrization rule, and the rule of BORN and JORDAN can be obtained from (26) by taking

$$f = 1, \quad \cos\frac{1}{2}\vartheta\tau\hbar, \quad \frac{\sin\frac{1}{2}\vartheta\tau\hbar}{\frac{1}{2}\vartheta\tau\hbar}$$

respectively. The choice of

$$f = \frac{\sin\frac{1}{2}\vartheta\tau\hbar}{\frac{1}{2}\vartheta\tau\hbar}$$

when substituted in (15) yields the distribution function

$$F(q, p) = \frac{2}{\hbar} \frac{1}{4\pi^2} \iiint \frac{e^{-i\tau p - i\vartheta q + i\vartheta u}}{\vartheta\tau} \sin\tfrac{1}{2}\vartheta\tau\hbar \cdot \psi^*(u - \tfrac{1}{2}\tau\hbar)\, \psi(u + \tfrac{1}{2}\tau\hbar)\, d\vartheta\, d\tau\, du. \tag{30}$$

Having at our disposal the set of all possible correspondence rules, it is natural to ask whether any one of them can be applied in a consistent manner. By this we mean the following: In quantum mechanics the operator which represents a function, say K, of the operator $G(\boldsymbol{q}, \boldsymbol{p})$ is $K(G(\boldsymbol{q}, \boldsymbol{p}))$. Classically the observable which represents a function of $g(q, p)$ is also $K(g(q, p))$. Now, is it possible to find a rule of correspondence such that the same rule can be used not only to obtain $G(\boldsymbol{q}, \boldsymbol{p})$ from $g(q, p)$ but also $K(G(\boldsymbol{q}, \boldsymbol{p}))$ from $K(g(q, p))$? The following discussion will show that the answer to this question is no.

The question is tantamount to asking: can quantum mechanics be formulated as a normal stochastic theory? The fact that Eq. (15) generates all distribution functions which satisfy (10) and (11) leads us to consider the possibility of formulating quantum mechanics as a classical stochastic theory. If this could be accomplished then some of the "paradoxical", or better, novel statements of quantum mechanics could be cast in more conventional language with the possibility of their resolution in terms of classical concepts. Let us list the requirements which the distribution function, $F(q, p, t)$, must satisfy if quantum mechanics is to be cast as a classical probability theory.

They must include the conditions 2, 3 and 4 which served as the basis for our discussion of measurement theory, but they also require that the formation shall yield the correct quantum mechanical expectation values for all observables. We therefore make the following more stringent demands.

a) Since the distribution F should be a probability function, it must be non-negative definite for all values of q and p.

b) The distribution function must yield the correct quantum mechanical marginal distribution functions, as before.

$$\int F(q, p)\, dp = |\psi(q)|^2,$$
$$\int F(q, p)\, dq = |\varphi(p)|^2.$$

As these were the conditions used to built up the set of functions given by Eq. (15), we automatically know that any F does indeed yield the marginal distributions, provided condition (16) is satisfied by $f(\vartheta, \tau)$.

c) The expectation values of observables obtained through phase space integration, using F, must give the same results as would be obtained by use of the quantum mechanical methods. That is, if $g(q, p)$ is the classical function to which corresponds the quantum mechanical operator $G(\mathbf{q}, \mathbf{p})$, then we require that

$$\text{c}_1) \qquad \langle\psi|\, G(\mathbf{q}, \mathbf{p})\, |\psi\rangle = \iint g(q, p)\, F(q, p)\, dq\, dp$$

and also that, for any function K,

$$\text{c}_2) \qquad \langle\psi|\, K(G(\mathbf{q}, \mathbf{p}))\, |\psi\rangle = \iint K(g(q, p))\, F(q, p)\, dq\, dp.$$

Condition c_1) can always be satisfied. But once an F is chosen to satisfy c_1), this same F cannot be used to calculate the expectation value of $K(g(q, p))$. Hence c_2) cannot in general be satisfied of c_1) is true. The proof can be found in (11).

These results also answer the question we raised. For it is clear that if it were possible to find an f such that

$$g(q, p) \to G(\mathbf{q}, \mathbf{p})$$

and

$$K(g(q, p)) \to K(G(\mathbf{q}, \mathbf{p}))$$

then c_1) and c_2) would be compatible for the same f.

We believe that these considerations, and the proof to which we have referred, definitively settle a question which has been asked and answered inconclusively in many different ways. Our answer does not, however, imply that no joint probability distribution exists which is compatible with the uncertainty principle for position and momentum. For example, F_0 of Eq. (14), meets all the requirements of a probability distribution, and yet the uncertainty principle does follow from it [8].

References

[1] ALBERTSEN, J.: Phys. Rev. **129**, 940 (1963). — ARAKI, H., and M. YANASEE: Phys. Rev. **120**, 622 (1961). — DURAND, III. L.: Phil. Sci. **27**, (2) (1960). — HEISENBERG, W.: La nature dans la physique contemporaine. Paris: Gallimard 1962; — Trad. de „Das Naturbild der heutigen Physik". Hamburg: Rowolt 1955. — JAUCH, J. M.: Helv. Phys. Acta **33**, 711 (1960). — JAUCH, J. M., and G. PIRON: Helv. Phys. Acta **36**, 827 (1963). — LANDÉ, A.: Phys. Rev. **108**, 891 (1957). — SCHILPP, P. (ed.): Albert Einstein: Philosopher-Scientist. Evanston: Library of Living Philosophers 1949. — SHIMONY, A.: Am. J. Phys. **31**, 755 (1963). — TISZA, L.: Revs. Mod. Phys. **35**, 151 (1963). — WIGNER, E. P.: Am. J. Phys. **31**, 6 (1963). An excellent, fully documented historical treatment of views here discussed is found in MAX JAMMER, The conceptual development of the quantum theory. New York: McGraw-Hill Book Co. 1966. — The contemporary scene is well surveyed by BERNARD D'ESPAGNAT, Conceptions de la Physique Contemporaine. Paris: Herman 1965.
[2] HEISENBERG, W.: Physics and philosophy. Harper 1958.
[3] MARGENAU, H.: The nature of physical reality, p. 171 et seq. New York: McGraw-Hill Book Co. 1950.
[4] — Phil. Sci. **30**, 1, 138 (1963).
[5] — The nature of physical reality. Ref. 3, chap. 13. Further discussion of the objectivity problem may be found. In: H. MARGENAU and J. PARK: Objectivity in quantum mechanics. In: M. BUNGE, ed.: Delaware Seminar in the Foundations of Physics. New York: Springer 1967.
[6] JEFFREYS, H.: Scientific inference. New York: McMillan 1931.
[7] MARGENAU, H.: Ann. of Phys. **23**, 469 (1963).
[8] —, and R. N. HILL: Progr. Theoret. Phys. (Kyoto) **26**, 722 (1961).
[9] WIGNER, E.: Phys. Rev. **40**, 749 (1932).
[10] COHEN, L.: J. Math. Phys. **7**, 781 (1966).
[11] — Thesis Yale University 1966, University Microfilm, Ann Arbor, Michigan.
[12] GROENEWOLD, H. J.: Physica **12**, 405 (1946).