PROBLEMS CONCERNING MEASUREMENT: A STUDY OF THE FOUNDATIONS OF QUANTUM THEORY

James L. Park

A Dissertation

Presented to the Faculty

of the Graduate School of Yale University

in Candidacy for the Degree

of

Doctor of Philosophy

1967

THESIS SUMMARY

Although the quantum era of natural philosophy dawned amid problems of interpretation, it is widely believed that conceptual difficulties associated with quantum physics were long ago resolved in Copenhagen. Thus modern physicists have concentrated primarily upon applications of the quantum algorithm and excluded from consideration deeper issues concerning its logical and epistemological structure. However, there has recently been a resurgence of interest in problems of this more philosophic kind; and the center of attention has been the concept of measurement. This dissertation explores problems of measurement in quantum theory largely from the vantage point of theoretical physics, but due emphasis is also placed on the philosophic aspects of any such basic inquiry. The work is divided into three self-contained parts.

Part I, "The Nature of Quantum States", carefully explores the relations among the statistical ensembles, systems, and states of quantum theory. By systematically contrasting the classical and quantal realizations of a general paradigm for a probabilistic physics, important distinctions are exposed both in statics and dynamics. It is concluded that the conceptual gulf between classical states and quantum "states" is wider than commonly assumed.

The overall purpose of Part II, "Quantum Theoretical Concepts of Measurement", is to clarify the physical meaning and epistemological status of the term measurement as used in quantum theory. After interpretive discussions contrasting the quantal concepts observable and ensemble with their classical ancestors along the lines of Margenau's latency theory, various popular ideas concerning the nature of quantum measurement are critically surveyed. A careful study of the quantum description of real experiments is then used to motivate a proposal that two distinct quantum theoretical measurement constructs should be recognized, both of which must be distinguished from the concept of preparation. The different epistemological roles of these concepts are compared and explained. It is concluded that the only possible type of "quantum measurement theory" is one of little metaphysical interest and that quantum measurement seems problematical only when viewed from an overly narrow classical perspective.

Part III, "Simultaneous Measurability in Quantum Theory", is a study of what is sometimes regarded as the conceptual heart of quantum theory, viz., the orthodox "physical" interpretation of noncommuting operators as representatives of incompatible (non-simultaneously-measurable) observables. It is demonstrated that the much quoted "principle" of incompatibility of noncommuting observables is simply false. The axiomatic root of all incompatibility arguments is then identified; and it is shown that, with a slight modification of the basic postulates which affects neither useful theorems nor practical calculations, quantum physics no longer entails illogical restrictions on measurability. Among the related topics touched upon is the problem of quantum joint probability distributions.

ACKNOWLEDGMENTS

It is a pleasure to express my sincere thanks to Professor Henry Margenau, who through his writings enabled me to discover as an undergraduate the challenging realm of physical foundations problems and who in numerous private discussions provided me vibrant intellectual stimulus for my graduate research.

I am also grateful to the National Aeronautics and Space Administration for the financial aid of a three-year fellowship.

And a special orchid must be tendered to my wife Marion, who valiantly typed not only this final Brobdingnagian thesis but also the penultimate--and antepenultimate--versions of it.

I. THE NATURE OF QUANTUM STATES

ABSTRACT

As a prerequisite to any meaningful understanding of the quantum measurement process, Part I carefully explores the relations among the statistical ensembles, systems, and states (pure and mixed) of quantum theory. By systematically contrasting the classical and quantal realizations of a general paradigm for a probabilistic physics, important distinctions are exposed both in statics and dynamics. Included are observations concerning the intrinsic ambiguity of the quantum state concept and the peculiarly quantum property of dynamic indivisibility. It is concluded that the conceptual gulf between classical states and quantum "states" is wider than commonly assumed.

1. Introduction

The dominant theme of the quantum theory--though many textbooks do not sufficiently emphasize it--is that all its causal statements are probabilistic. In other words, the epistemic rule of correspondence[1] which provides the empirical meaning of quantum theoretical states involves probabilistic concepts in an essential way. This hallmark of quantum theory must be constantly borne in mind if the physical significance of the theory is to be understood at all. Moreover, it is of fundamental importance to recognize that the probabilistic rule alone is insufficient to link abstract states with empirical experience. Indeed pure probability theory is itself a formidable collection of abstruse constructs requiring further rules of correspondence of its own. The situation is wholly analogous to that prevailing in geometry, where pure geometry is converted to physical geometry by appending familiar operational definitions involving straightedge and compasses. In the case of probability theory, the required connections are implicit in established practices of experimental science. Especially to be noted is the well known identification of physical probability as relative frequency* in a statistical ensemble. Accordingly, if quantum theory is to be understood as a physical science, it is absolutely essential to remember that its primary connection to the empirical world is through statistical ensembles. To ignore the statistical aspect of quantum theory is to dismiss much of its relevance to actual physical experience.

Nevertheless, it has proved intuitively useful in ordinary quantum theoretical applications to think of the state vector (or its wave function

---

*The appropriateness of the frequency definition of probability is an old philosophic problem, but we are not concerned with it here. We simply accept uncritically the fact that in all scientific applications of probability theory, the construct probability is epistemically linked to statistics.[2]

representative) as belonging to a single system at a single time in the same way that states belong to individual systems in classical mechanics. Thus the jargon of modern physics easily induces one to regard the phrase "an electron in state ψ" as merely the quantum analogue to the classical expression "an electron in state (q_0, p_0)", in spite of the fact that the former refers physically to statistics of measurement results upon an ensemble of identically prepared electrons* whereas the latter just means that a single (classical) electron has position q_0 and momentum p_0. Superficially, this common phraseology seems innocent enough; indeed one might be disposed to think that a theoretician could use it unreservedly without contradiction so long as he remembered to switch to the correct statistical meaning of ψ at the conclusion of his arguments and calculations.

It is the purpose of the present study to expose certain logical weaknesses inherent in the drawing of structural parallels between classical and quantum physics and hence to demonstrate that a quantum theorist may not in every context relate state vectors to single systems in the classical manner described above. The linguistic extension of ψ from its role in describing ensembles to its further function as the state of a single system has given birth to monumental barriers to the understanding of quantum theory as a rational branch of natural philosophy. Problems connected with the general theory of measurement--the nature of quantum measurement, wave packet reduction, concepts of compatibility and simultaneous measurement--are especially aggravated by this popular convention that the state of an individual system is represented by ψ.

Thus in our opinion the material to be surveyed below forms the

---

*The ensemble may consist of many independent electrons at one time, one electron sequentially measured and reprepared, or any combination of these two extremes.[3]

essential prelude to any serious study of the basic philosophic issues associated with quantum theory. We shall attempt to show elsewhere[4] that problems concerning measurement in quantum physics can be sharpened, and sometimes resolved, by according proper attention to those basic physical characteristics of quantum states with which the present essay deals.

2. The Minimal Axiomatic Structure of a Probabilistic Physics

Rather than devising in the customary way artificial verbal analogies between the constructs of classical and quantal physics (e.g., the aforementioned state (q_0, p_0) and "state" ψ), we shall contrast the two kinds of theories in a manner which reveals formal differences as readily as the standard comparisons indicate formal similarities. To facilitate such a comparison of classical and quantal statistical physics, it is useful to consider an abstract paradigm theory representative of a probabilistic physical theory in general. As in all physical theories, the primitive idea of this prototype theory is the study of the numerical results $\{a_r\}$, $\{b_s\}$,...of measurements of observables A, B,...performed on a physical system S. It is the goal of the theory to incorporate these measurement results into a causal framework, i.e., information about present measurement results should determine similar information about future measurement results. The defining property of the present model theory is the fact that this information has a form related to the probabilities of the measurement results. Hence what is actually studied is the statistics of the results of measurements on the member systems of ensembles of identically prepared replicas of the physical system of interest S.* In the theory, an ensemble is characterized at a given time by the

---

*Cf. fn. on p.2.

arithmetical mean values (expectation values) of measurement results for the various observables, i.e., by the "functional" $\langle a \rangle$ whose "domain" is the set of observables. In such a theory the causal nexus links the values of $\langle a \rangle$ at different times for every a.

The essential axioms which typify a statistical physical theory may be summarized as follows:

P1: Mathematical objects A, B,...correspond to observables $a, B, \ldots$

P2: For every ensemble of identically prepared systems there exists a functional $m_t(A)$ such that for every pair (A, a),

$$m_t(A) = \langle a \rangle_t ,$$

the arithmetic mean of the results of a-measurements performed at time t relative to the preparation of each member system.

P3: For every kind of physical system (i.e., system of interest plus its relevant environment) there exists a causal law relating the forms of the mean value functional at different times.

Later more definite content will be ascribed to these rather vague statements. First, however, we must review an important scheme for the classification of statistical ensembles especially emphasized by von Neumann.[5] This classification hinges on the concept of ensemble homogeneity, a property which ultimately depends on the mathematical character of measurement statistics associated with the ensemble. In particular, it is always possible to conceive of many subdivisions of a given ensemble into subensembles*; the homogeneity of the original ensemble is determined by comparing the statistical characteristics of such subensembles. This process of subdivision is of course a <u>mental</u> operation based on statistics of measurement results; indeed it should always be remembered that the ensemble

*These subensembles must themselves be bonafide statistical ensembles, i.e., they must each contain an "effectively infinite" number of systems so that the frequency definition of probability may still be used.

concept includes even the case of a single system in a temporal sequence of measurements and repreparations. Most ensembles are mixed, in the sense that the subensembles into which they may be grouped are statistically distinct; there exist, however, completely homogeneous, or pure, ensembles for which every subdivision yields subensembles all statistically identical to the original. Mathematically, a mixed ensemble is characterized by a mean value functional $m(A)$ such that there exist distinct functionals $m_1(A), m_2(A)$ in terms of which $m(A)$ may be expressed as follows:

$$m(A) = w_1 m_1(A) + w_2 m_2(A), \quad w_1, w_2 > 0,$$

where w_1, w_2 denote the respective fractions of the systems in the original ensemble present in subensembles 1 and 2. (Clearly, $w_1 + w_2 = 1$.) If, on the other hand, for a given $m(A)$ there do not exist distinct $m_1(A), m_2(A)$ such that $m(A) = w_1 m_1(A) + w_2 m_2(A)$, the ensemble characterized by that $m(A)$ is of the pure, or homogeneous type. As we shall see later, it is the latter type of ensemble which, due to its maximal order and uniformity, may in some cases provide a means to extract from an initially probabilistic theory (one dealing basically with ensembles and statistics) a plausible state concept applicable to single systems in the classical sense, i.e., a state concept such that any physical system considered at a given instant may be regarded without contradiction as having a definite state at that instant.

To exhibit a simple realization of the general paradigm above and to exemplify the concept of ensemble homogeneity, it is instructive to examine classical statistical mechanics from an unusual perspective which disregards the original purpose of that classical theory as a mechanical explanation of thermodynamics. That is, we are not interested in the Gibbsian imaginary ensemble of replicas and do not adopt the usual postulate which connects averages over that ensemble with thermodynamic parameters

associated with the single system of interest. Instead, we consider a real ensemble (either an aggregate of identically prepared systems upon which measurements are performed or an alternating temporal sequence of preparations and measurements upon a single system, or a combination of both) and the collectives of measurement results obtained by real measurement operations upon its member systems. Moreover, assume that nothing is known about Newtonian properties or states of single systems. From this point of view, the sole purpose of classical statistical mechanics is to regularize within a causal framework purely probabilistic information about measurement results; hence the theory is a realization of the above paradigm and may be summarized as follows:

P1C: Functions $A(q,p)$, $B(q,p)$,...whose domain is phase space correspond to observables $\mathcal{A}$, $\mathcal{B}$,...(Phase space has for coordinates the position (q), and momentum (p) components of the system of interest.)

P2C: For every ensemble of identically prepared systems there exists a mean value functional $m_t[A(q,p)]$. Every m_t may be expressed in terms of a corresponding function $\rho_t(q,p)$ (the density-of-phase) as follows:

$$m_t(A) = \int dq \int dp\, \rho_t(q,p) A(q,p).$$

(It can be shown that $\rho \geq 0$, $\rho \not\equiv 0$, and $\int dq \int dp\, \rho = 1$.)

P3C: For every kind of physical system, there exists a function $H(q,p)$ (the Hamiltonian) which determines the causal evolution of m_t via the following law (Liouville's equation):

$$\frac{\partial \rho_t}{\partial t} = \sum \left(\frac{\partial H}{\partial q} \frac{\partial \rho}{\partial p} - \frac{\partial H}{\partial p} \frac{\partial \rho}{\partial q} \right) \equiv \{H, \rho\},$$

where $\sum$ denotes summation over all conjugate pairs (q,p).

The pure and mixed ensembles of this theory are easily identified by combining P2C with the general definitions of pure and mixed. In terms of

the density-of-phase function, an ensemble characterized by $\rho(q,p)$ is pure if there do not exist distinct (nonzero) functions $\rho^{(1)}(q,p)$, $\rho^{(2)}(q,p)$ such that

$$\rho(q,p) = w_1\rho^{(1)}(q,p) + w_2\rho^{(2)}(q,p), \quad w_1, w_2 > 0.$$

Now, it is obvious from analytic geometry considerations that any function $\rho(q,p)$ whose support (i.e., that part of the domain for which the function is nonzero) may be divided into two regions R_1 and R_2 can indeed be written as a linear combination of distinct functions $\rho^{(1)}$, $\rho^{(2)}$ by letting $\rho^{(1)}(\rho^{(2)})$ be proportional to ρ for points (q,p) in $R_1(R_2)$ and zero for points in $R_2(R_1)$. Thus the support of a pure $\rho(q,p)$ must be indivisible, i.e., a single point of phase space, say (q_0,p_0). The generalized function which satisfies this requirement plus the normalization condition $\int dq \int dp\, \rho(q,p) = 1$ is just the Dirac delta:

$$\rho(q,p) = \delta(q-q_0)\,\delta(p-p_0).$$

This is therefore the general form of ρ for the pure ensembles of classical statistical mechanics; particularly to be noted is the correspondence between these pure ensembles and the points (q_0,p_0) of phase space, a correspondence which will play a central role in subsequent development of a state concept applicable to a single system at a single time. Finally, the mixed ensembles of classical statistical mechanics are simply those represented by density functions not in the above pure form.

3. Augmentation of the Model Theory to Include a State Concept

The minimal axiomatic structure (P1, P2, P3) outlined above for an intrinsically probabilistic physical theory does not mention the concept of state. The omission was deliberate, for the usefulness of the state concept in physics lies in its reference to individual systems and its participation in the scheme of general causality (initial state of a system implies final state of that system via physical laws). Within the

purely statistical model alone, such states play no role. There is, however, the possibility that they might be added somehow to that paradigm, provided the supplementation yields a consistent logical structure. It will be demonstrated below that the concept of ensemble homogeneity provides both a suitable means for generating such a supplemental postulate and a touchstone to check for ambiguities or inconsistencies within a theory thus modified.

To motivate this general supplementation procedure, let us first reconsider classical statistical mechanics in the special form given above, where the axiomatic structure (P1, P2, P3) does not involve a priori any state concept for single systems. It is, however, well known that such a concept can indeed be consistently appended to this structure. In fact, the usual theoretical development of statistical mechanics begins with the classical states (q_0, p_0) and the Newtonian laws which govern their causal evolution; $\rho(q,p)$ is then defined so that $\int_{q_1}^{q_2} dq \int_{p_1}^{p_2} dp\, \rho(q,p)$ equals that fraction of an ensemble whose systems have classical states (q,p) in the phase space region $\{(q,p) \mid q \in (q_1, q_2), p \in (p_1, p_2)\}$, and Liouville's equation (P3C) is derived by applying classical mechanics to each member system of the ensemble.

In the present context, we are interested in the opposite procedure, viz., extracting from an originally statistical theory a consistent state concept for single systems. Thus in statistical mechanics the $\rho(q,p)$ is to be thought of initially only as characterizing an ensemble in toto and not as being formed by counting numbers of elements "really in" various classical states (q,p). The latter interpretation is what we want to "discover" from an analysis of measurement statistics associated with the ensembles. The fundamental question to be posed, then, is the following: Can an ensemble, described by $\rho(q,p)$, be construed <u>unambiguously</u> as

consisting of elements to which some meaningful state specification may be assigned individually?

For classical statistical mechanics, the affirmative reply comes about as follows. Consider first the pure ensemble. Since, by definition, every division, or selection, produces subensembles statistically identical to each other and to the original, the pure ensemble provides the natural starting point for any additional postulate about individual states. Indeed, the homogeneity under division exhibited by the pure ensemble $\rho(q,p)=\delta(q-q_0)\delta(p-p_0)$ strongly suggests a supplemental postulate which assigns to each member system of such an ensemble a state (q_0,p_0), which may be thought of as possessed by the individual system. In classical statistical mechanics, the fact that observables are represented by functions A(q,p) strengthens this suggestion; for thoughtful analysis clearly indicates that if for every A(q,p), $m(A)=A(q_0,p_0)$ (as is the case when $\rho=\delta(q-q_0)\delta(p-p_0)$), the only reasonable definition for the possessed state of an individual system is the traditional phase space point (q_0,p_0) (or an equivalent specification in terms of the associated values of other phase functions).

This observation is, however, not yet sufficient to warrant inclusion of the proposed supplemental postulate; mixed ensembles have yet to be considered. Since mixed ensembles are characterized by reducibility to distinct subensembles and pure ensembles permit the unambiguous assignment of individual states, the fundamental question posed earlier may be replaced by another: Is the reduction of a mixed ensemble to a set of pure subensembles unique? If so, any ensemble may be consistently interpreted as a collection of systems individually described by definite states; if not, the proposed assignment of states to single systems is ambiguous and therefore physically meaningless.

In classical statistical mechanics, this reduction to pure subensembles is represented mathematically by expressing a general $\rho(q,p)$ as a "sum" of $\delta(q-q_0)\,\delta(p-p_0)$:

$$\rho(q,p) = \int dq_0 \int dp_0\, C(q_0,p_0)\,\delta(q-q_0)\,\delta(p-p_0).$$

Necessarily, $C(q_0,p_0) = \rho(q_0,p_0)$, a unique solution. Therefore, every member of a classical ensemble may be unambiguously assigned a classical state (q_0,p_0) at any given time.

As noted above, the present discussion of statistical mechanics has proceeded retrograde to the traditional development. The ensemble and measurement statistics have been regarded as primitive, while the notion of state has been sought through an analysis of the overall theoretical structure of the ensemble, in contrast to the standard procedure of forming ensembles from systems in mechanical states defined from the beginning. What has been done might be described rather crudely as the extraction of classical mechanics from statistical mechanics, instead of vice versa. More accurately, a study of the structure of classical statistics has shown that the state concept of classical mechanics may be "derived" by supplementing statistical mechanics with an additional state specification postulate, as opposed to the historic procedure of constructing statistical mechanics as the union of classical mechanics and statistics. To justify fully the identification of phase space points (q_0,p_0) as states, it is of course also necessary to prove that pure ensembles evolve into pure ensembles through Liouville's equation. This will be done subsequently in connection with a systematic comparison of classical and quantal dynamics (sec. 7). To summarize, the concept state (of a single system at a single time) may be unambiguously attached to statistical mechanics as a fourth postulate:

P4C: Every system is always in a state represented by some point (q_0,p_0) of phase space.

The preceding discussion of classical statistical physics was given to illustrate the possibility of and the method for extracting the individual state concept from a theory initially concerned only with ensembles. As explained at the outset, the motivation for such an analysis comes from quantum theory, a discipline for which any sensible discussion of the state concept must proceed from the root notion of statistical ensemble owing to the intrinsic probabilistic nature of the theory. Thus the proper approach to the meaning of states in quantum theory is along the lines of the foregoing "reverse" development of statistical mechanics.

To study critically the nature of quantum states, we must therefore ascertain whether or not quantum theory, like statistical mechanics, is a realization of the general model given earlier (P1, P2, P3) augmented by a fourth postulate of the following form:

P4: Every system is always in a state represented by an element belonging to a set of mathematical objects which correspond to the pure ensembles of the theory.

The logical admissibility of P4 to any given theory always depends on the exact content of P1-P3. In particular, the state concept will be unambiguous if and only if the resolution of a general mixed ensemble into pure subensembles is unique. Moreover, if the state identification is to be physically meaningful, initially pure ensembles must retain their homogeneity under causal evolution.

4. Is Quantum Theory a Realization of the Augmented Paradigm?

At first glance, the augmented model seems tailor made for quantum theory; this should be no surprise, since it was obviously inspired by the standard axiomatic pattern of quantum mechanics texts. However, to avoid jumping to conclusions, a stepwise analysis of the strictness of analogy

between quantum theory and the general paradigm is in order. Accordingly, the postulates of quantum theory will be scrutinized one by one and with an alertness to their consistency and interdependence.

P1Q: The linear Hermitean operators A, B,...on Hilbert space which have complete orthonormal sets of eigenvectors correspond to physical observables $a, B, \ldots$ If operator A corresponds to observable a, then the operator $\mathcal{F}(A)$ corresponds to observable $\mathcal{F}(a)$, where $\mathcal{F}$ is a function.

P2Q: For every ensemble of identically prepared systems there exists a real linear mean value functional $m_t(A)$ defined on the Hermitean operators.

Superficially, P1Q and P2Q seem almost too vague to have significant consequences. The vagueness is an illusion. In fact, P1Q and P2Q imply many well known features of quantum theory. In particular, it follows* that every $m_t(A)$ may be expressed in terms of a corresponding operator ρ_t (the density operator) as follows: $m_t(A) = Tr(\rho_t A)$. (It can be shown that ρ is Hermitean, positive semidefinite, and that $Tr\rho = 1$.) Moreover, P1Q and P2Q imply that the probability for an a-measurement to yield a result other than an A-eigenvalue is zero.

There are several ways to express the dynamical postulate of quantum theory; perhaps the following version is best in the present context since we are specifically interested in contrasting quantum theory with classical statistical mechanics.

P3Q: For every kind of physical system, there exists an Hermitean

---

*For a full discussion of this and other theorems, see ref. . Incidentally, it should be clear that the various sets of postulates under consideration presently and in preceding sections are certainly not complete in any rigorous logical sense. Rather attention is purposely directed to what might be termed "physical" postulates to the exclusion of such necessary supporting axioms as those of function and functional analysis, linear algebra, probability calculus, statistics, etc.

operator H (the Hamiltonian) which determines the causal evolution of m_t via the following law:

$$i\hbar \frac{\partial \rho_t}{\partial t} = H\rho - \rho H \equiv [H, \rho].$$

P1Q, P2Q, and P3Q make up the axiomatic core of quantum theory; all the statistical results of quantum physics are derivable within this framework. Thus while a rather Procrustean reversal of foundations and consequences was required to force classical statistical mechanics into our model, quantum theory fits naturally. What seemed to be a "reverse" development in the classical case is the only sensible one in quantum physics, for there is no nonprobabilistic "mechanics" applicable to individual quantum systems. Since probabilities are present in quantum mechanics from the very outset, statistical ensembles are the physical objects of study from the very beginning.

Nevertheless, it is conceivable that some abstract object related to the pure ensemble might be theoretically attached to every quantum system as its state by an analysis parallel to that which led to P4C in the classical case. In short, quantum theory might be a realization of the augmented model (P1, P2, P3, and P4). To investigate this possibility, the pure ensembles of the theory must be found. In terms of the density operator ρ, an ensemble characterized by ρ is pure if there do not exist distinct (nonzero) positive semidefinite Hermitean operators $\rho^{(1)}, \rho^{(2)}$ such that

$$\rho = w_1 \rho^{(1)} + w_2 \rho^{(2)}, \quad w_1, w_2 > 0$$

The desired identification of the pure ensembles of quantum theory is accomplished by the following theorem due to von Neumann:

(H) ρ is pure if and only if $\rho = P_\psi$, where P_ψ is a projection operator onto the Hilbert vector ψ.

Because the proof[6] is seldom repeated and may be unfamiliar, we shall digress briefly to outline the argument.

(H_1) If ρ is pure, $\rho = P_\psi$.

Proof: Let χ be a unit vector such that $\rho\chi \neq 0$ and define Hermitean operators $\rho^{(1)}, \rho^{(2)}$ as follows:

$$w_1\rho^{(1)} \equiv \frac{\rho P_\chi \rho}{\langle \chi, \rho\chi\rangle}, \quad w_2\rho^{(2)} = \rho - w_1\rho^{(1)},$$

where $w_1, w_2 \geq 0$, $w_1 + w_2 = 1$, $\mathrm{Tr}\,\rho^{(1)} = \mathrm{Tr}\,\rho^{(2)} = 1$ ($\langle , \rangle$ denotes the scalar product.) Since ρ is positive semidefinite, $\langle \chi, \rho\chi\rangle > 0$; thus

$$w_1\langle \phi, \rho^{(1)}\phi\rangle = \frac{\langle \phi, \rho P_\chi \rho\phi\rangle}{\langle \chi, \rho\chi\rangle} = \frac{|\langle \phi, \rho\chi\rangle|^2}{\langle \chi, \rho\chi\rangle} \geq 0,$$

for every vector ϕ, i.e., $\rho^{(1)}$ is positive semidefinite.

Now, every definite Hermitean operator D satisfies the following inequality*:

$$|\langle \phi_n, D\phi_m\rangle|^2 \leq \langle \phi_n, D\phi_n\rangle\langle \phi_m, D\phi_m\rangle, \text{ for every } \phi_n, \phi_m.$$

Hence, for every ϕ,

$$w_2\langle \phi, \rho^{(2)}\phi\rangle = \frac{\langle \phi, \rho\phi\rangle\langle \chi, \rho\chi\rangle - |\langle \phi, \rho\chi\rangle|^2}{\langle \chi, \rho\chi\rangle} \geq 0$$

i.e., $\rho^{(2)}$ is positive semidefinite.

Thus $\rho^{(1)}$ and $\rho^{(2)}$ are admissible density operators, and the expression $\rho = w_1\rho^{(1)} + w_2\rho^{(2)}$ represents a subdivision of the original ensemble (ρ) into subensembles characterized by $\rho^{(1)}$ and $\rho^{(2)}$, w_1 and w_2 being the respective fractions of the whole ensemble present in the two subensembles.

Since ρ is pure, by definition $\rho^{(1)}$ cannot differ from $\rho^{(2)}$; hence for every ϕ,

$$\rho\phi = \rho^{(1)}\phi = \frac{1}{w_1}\frac{\rho P_\chi \rho\phi}{\langle \chi, \rho\chi\rangle} = \frac{1}{w_1}\frac{\langle \rho\chi, \rho\chi\rangle P_{\rho\chi}\phi}{\langle \chi, \rho\chi\rangle}.$$

Define $\psi = \rho\chi$; then

$$\rho\phi = \frac{1}{w_1}\frac{\langle \psi, \psi\rangle}{\langle \chi, \psi\rangle}P_\psi\phi, \text{ for every } \phi. \text{ Therefore,}$$

$$\rho = \left(\frac{1}{w_1}\frac{\langle \psi, \psi\rangle}{\langle \chi, \psi\rangle}\right)P_\psi \text{ or since } \mathrm{Tr}\,\rho = 1, \ \rho = P_\psi.$$

(H_2) If $\rho = P_\psi$, ρ represents a pure ensemble.

---

*For a proof of this lemma, see Ref. 5, p. 101.

Proof: Let η be a vector such that $\langle \psi, \eta \rangle = 0$. Consider an expansion $\rho = w_1 \rho^{(1)} + w_2 \rho^{(2)}$, $w_1, w_2 \geq 0$, $w_1 + w_2 = 1$, where $\rho^{(1)}, \rho^{(2)}$ are density operators. Since $\rho^{(1)}, \rho^{(2)}$ are positive semi-definite,

$$0 \leq w_1 \langle \eta, \rho^{(1)} \eta \rangle \leq w_1 \langle \eta, \rho^{(1)} \eta \rangle + w_2 \langle \eta, \rho^{(2)} \eta \rangle$$
$$= \langle \eta, \rho \eta \rangle = \langle \eta, P_\psi \eta \rangle = \frac{\langle \psi, \eta \rangle \langle \eta, \psi \rangle}{\langle \psi, \psi \rangle} = 0,$$

i.e., $\langle \eta, \rho^{(1)} \eta \rangle = 0$.

Applying to $\rho^{(1)}$ the inequality introduced above, viz., $|\langle \phi_n, D \phi_m \rangle|^2 \leq \langle \phi_n, D \phi_n \rangle \langle \phi_m, D \phi_m \rangle$, D Hermitian definite, we obtain $|\langle \theta, \rho^{(1)} \eta \rangle|^2 \leq \langle \theta, \rho^{(1)} \theta \rangle \langle \eta, \rho^{(1)} \eta \rangle = 0$, for each θ. Hence $\langle \theta, \rho^{(1)} \eta \rangle = 0$ for each θ.

Now, any vector η which satisfies $\langle \psi, \eta \rangle = 0$ may be expressed in the form $(1 - P_\psi)\gamma$. Thus the preceding equation may be written as follows: for each θ, γ, $\langle \theta, \rho^{(1)}(1 - P_\psi)\gamma \rangle = 0$. This implies that $\rho^{(1)} = \rho^{(1)} P_\psi$. From matrix considerations it is then fairly easy to see that $\rho^{(1)} = P_\psi = \rho$; hence by definition ρ represents a pure ensemble. Q.E.D.

To summarize, the foregoing theorem establishes a correspondence between the pure ensembles of quantum theory and the points (vectors*) of Hilbert space, a correspondence strikingly reminiscent of that between classical pure ensembles and the points of phase space. Thus once again it seems natural to take advantage of the full homogeneity of the pure ensemble--the defining characteristic that every subensemble is indistinguishable from the original by any measurement statistics--and assign a state to each individual system. In the present case, the vector ψ is

---

*Strictly speaking, since $P_\psi = P_{a\psi}$ for any complex number a, the pure ensembles correspond to rays rather than specific vectors, where a ray is a collection of all vectors expressible in the form $a\psi$, i.e., all vectors "pointing in the same direction."

clearly the appropriate representative for such a state. The stage is thus set for the seemingly harmless and perhaps intuitively useful jargon which makes statements of this type: "Suppose an electron is in the state ψ." Formally speaking, it seems reasonable to supplement P1Q, P2Q, P3Q by a P4Q which would assert that every quantum system is always in some state

It would, however, be premature at this point to make such an addition to the quantal framework; first we must determine whether or not the proposed state specification scheme makes sense, i.e., whether or not it is ambiguous. As explained earlier, unambiguous individual state specification via the pure ensemble is possible only if the resolution of a general ensemble into pure subensembles is unique. Just as in classical statistical mechanics, this is a mathematical question strongly dependent on the precise content of P1, P2, P3. Since the pure form of the density operator has been identified as the projection P_ψ, what is in question is the uniqueness of the sum,

$$\rho = \sum_k w_k P_{\psi_k}.$$

The surprising answer, which quantum theorists must face with all its ramifications, is the negative one. A general quantum ensemble can be subdivided in an infinite variety of ways into pure subensembles. As a result, the analogy between classical and quantal statistics breaks down. Quantum physics is not a realization of the foregoing augmented paradigm. There can be no "P4Q". Especially noteworthy is the consequence that it is generally improper to assign quantum state vectors to individual systems. To do so ultimately leads only to paradoxes, as will be illustrated below.

If all theoretical considerations in quantum theory could be carried out using only pure ensembles, the rather natural assignment of state vectors to single systems would be quite inconsequential. Thus, for instance, no bewildering paradoxes arise in the traditional applications of quantum

mechanics wherein individual state specification is a commonplace notion. However, mixtures cannot always be circumvented; in particular, even if a set of given systems constitute a pure ensemble, the ensemble formed from specified subsystems of these is in general mixed*. Hence the most general quantum ensembles--the mixtures--cannot be ignored.

To demonstrate that the assignment of state vectors to single quantum systems is essentially ambiguous and therefore improper, it suffices to give a simple illustrative counterexample.

Consider an ensemble of "spins", i.e., the associated Hilbert space is the familiar two-dimensional spinor space. Let $\alpha, \beta, \delta, \gamma$ denote eigenvectors of Pauli spin operators as follows:

$$\sigma_z \alpha = \alpha \,, \quad \sigma_z \beta = -\beta \,, \quad \sigma_x \delta = \delta \,, \quad \sigma_x \gamma = -\gamma \,.$$

Let the statistical operator describing the ensemble at some given time be $\rho^{(1)} = \frac{3}{4} P_\alpha + \frac{1}{4} P_\beta$. It is obvious from the structure of $\rho^{(1)}$ that a proper selection will yield two subensembles divided as follows: one consisting of 3/4 of the original systems and characterized by α , another made up of the remaining $\frac{1}{4}$ and characterized by β . One is inclined to describe this character of the ensemble by the statement that 3/4 of the original systems are "in state α ", $\frac{1}{4}$ "in state β ". It will now be shown that such an extrapolation from ensemble to single system leads to a paradox.

Consider a second ensemble of "spins" characterized by the statistical operator $\rho^{(2)} = \frac{3}{8} P_\delta + \frac{5}{8} P_\eta$, where $\eta = \frac{1}{\sqrt{5}} (\delta + 2\gamma)$. This ensemble may be partitioned by selection into two pure subensembles: one made up of 3/8 of the original systems and characterized by δ , the second consisting of 5/8 of the initial ensemble and characterized by η .

---

*Cf. sections 5 and 8.

Following the natural procedure for state specification, the ensemble might then be described as a collection of systems 3/8 of which are "in the state δ ", 5/8 "in the state η ".

It is useful at this stage to exhibit the matrix representative of the operator $\rho^{(2)}$ in the $\{\alpha,\beta\}$-representation. The matrix elements are easily calculated: using $\delta = \frac{1}{\sqrt{2}}(\alpha+\beta)$, $\gamma = \frac{1}{\sqrt{2}}(\alpha-\beta)$, it follows that

$$(P_\eta) = \begin{pmatrix} \langle\alpha, P_\eta \alpha\rangle & \langle\alpha, P_\eta \beta\rangle \\ \langle\beta, P_\eta \alpha\rangle & \langle\beta, P_\eta \beta\rangle \end{pmatrix} = \begin{pmatrix} 9/10 & -3/10 \\ -3/10 & 1/10 \end{pmatrix}; \quad (P_\delta) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix},$$

and $$(\rho^{(2)}) = \frac{5}{8}\begin{pmatrix} 9/10 & -3/10 \\ -3/10 & 1/10 \end{pmatrix} + \frac{3}{8}\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 3/4 & 0 \\ 0 & 1/4 \end{pmatrix}.$$

Hence, $$(\rho^{(2)}) = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

or $$\rho^{(2)} = \tfrac{3}{4} P_\alpha + \tfrac{1}{4} P_\beta .$$

Comparison now shows that $\rho^{(1)} = \rho^{(2)}$; i.e., the statistical operators characterizing the two spin ensembles considered are equal. Since the statistical operator completely describes all measurement results for an ensemble, the equality of $\rho^{(1)}$ and $\rho^{(2)}$ implies the <u>physical identity</u> of the ensembles they represent. Such an occurrence is by no means exceptional; this example is <u>not</u> a mathematical freak, or, as physicists sometimes say, a "pathological" case. Rather it illustrates a typical property of quantum mechanical ensembles, viz., that they are generally resolvable into <u>pure</u> subensembles in many ways, a property which was given above as the fundamental reason why quantum theory is not a realization of the augmented paradigm.

The basis for that conclusion may perhaps be clarified by further scrutinizing the present example: $\rho = \frac{3}{4} P_\alpha + \frac{1}{4} P_\beta = \frac{3}{8} P_\delta + \frac{5}{8} P_\eta$. From the structure of these resolutions of ρ into pure subensembles, it follows that there exist member systems which may belong either to a P_α-ensemble or a P_δ-ensemble. The following picturization of the ensembles

involved is helpful in understanding this. Let each member system be denoted by a box so that the ensemble may be pictured as a row of boxes.

The resolution $\rho = \frac{3}{4}P_\alpha + \frac{1}{4}P_\beta$ means that the boxes may be re-arranged and partitioned into two rows, one containing 3/4 of the original boxes and physically describable by P_α, the second containing the remaining $\frac{1}{4}$ and physically describable by P_β.

Re-arrangement must not be confused with physical interaction processes; what is considered here is the theoretical and conceptual structure of the ρ-ensemble.

Similarly, the resolution $\rho = \frac{3}{8}P_\delta + \frac{5}{8}P_\eta$ means that the boxes may also be re-arranged and partitioned into two different rows, one containing 3/8 of the original boxes and physically describable by P_δ, the second containing the remaining 5/8 and physically describable by P_η.

A comparative examination of the two arrangements makes intuitively clear the above assertion that there are systems in the ρ-ensemble which, without physical change, may belong to either a P_α-ensemble or a P_δ-ensemble. Now suppose for a moment the language of individual states is applied to these ensembles. Immediately a quantum monster is born: a <u>single</u> system concurrently "in" <u>two</u> states α and δ. To be explicit, the "state" of the system would have to be simultaneously an eigenvector of σ_z and σ_x, which is simply a mathematical impossibility. Thus the concept of individual quantum state is fraught with ambiguity and should therefore be avoided in serious philosophic inquiries concerning the nature of quantum theory.

5. <u>The Process of Selection: An Example</u>

The counterexample used in the preceding section was treated from an abstract viewpoint; in particular, ensembles were represented as collections

of "boxes" subject to "re-arrangement" and "partition" into subcollections. To demonstrate that the foregoing analysis does indeed have physical significance, a gedankenexperiment will now be outlined to illustrate the process of selection of pure subensembles and its inherent ambiguity.

Consider an ensemble the typical element of which is a pair of distinguishable, noninteracting atoms; let the initial density operator for this ensemble be $\rho_0 = P_\psi$, $\psi = \frac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 - \beta_1 \otimes \alpha_2)$, where α, β denote eigenvectors of σ_z, and vectors belonging to the spin spaces $\mathcal{H}_1, \mathcal{H}_2$ associated with atoms 1 and 2 are identified by subscripts 1, 2. Such an ensemble might be prepared by taking unstable molecules about which it is known with certainty that total $\vec{\sigma}^2$ and σ_z -measurements would both yield zero and that the disintegration process conserves spin. The pair of noninteracting atoms which results from such a disintegration would be a typical element of an ensemble described by $\psi = \frac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 - \beta_1 \otimes \alpha_2)$. For convenience let the "spins" move freely, i.e., assume $\rho_t = \rho_0 = P_\psi$ for all t* prior to the performance of certain operations to be specified below.

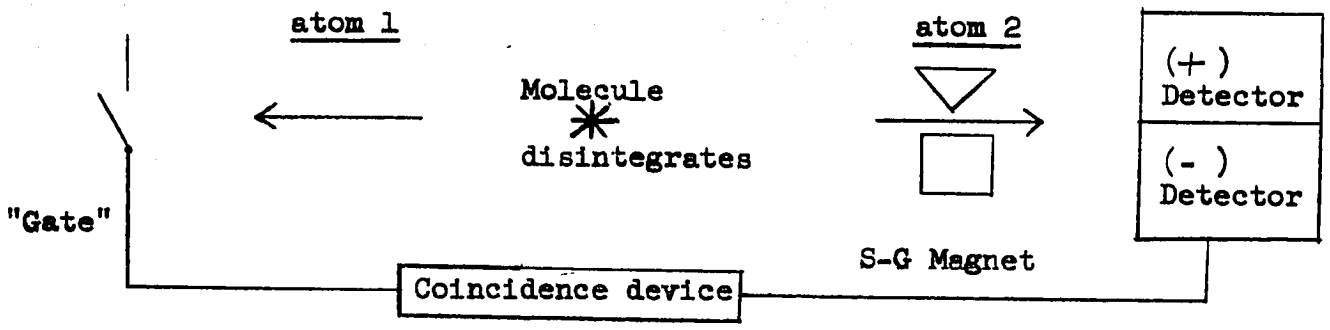
Now, consider the ensemble comprised of one atom from each pair, say atom 1. To determine the density operator for this ensemble, note that only Hermitean operators defined on $\mathcal{H}_1$ will be relevant when atom 1 is considered alone; i.e., we are now interested only in operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$ which have the form $A_1 \otimes 1$. The mean value functional for such operators has the following structure:

$$
\begin{aligned}
m(A_1 \otimes 1) &= \mathrm{Tr}(P_\psi A_1 \otimes 1) = \langle \psi, A_1 \otimes 1 \psi \rangle \\
&= \left\langle \tfrac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 - \beta_1 \otimes \alpha_2), \tfrac{1}{\sqrt{2}}(A_1\alpha_1 \otimes \beta_2 - A_1\beta_1 \otimes \alpha_2) \right\rangle \\
&= \tfrac{1}{2}\left(\langle \alpha_1, A_1\alpha_1 \rangle + \langle \beta_1, A_1\beta_1 \rangle\right) = \tfrac{1}{2}\mathrm{Tr}_1(A_1) \\
&= \tfrac{1}{2}\mathrm{Tr}_1(1A_1) = \mathrm{Tr}_1\left[\tfrac{1}{2}(P_{\alpha_1} + P_{\beta_1})A_1\right].
\end{aligned}
$$

*For simplicity, configuration space parts of ψ have been suppressed. The spinor structure of ψ is unchanged during free evolution.

Thus the density operator $\rho_1 = \frac{1}{2}P_{\alpha_1} + \frac{1}{2}P_{\beta_1}$ describes the ensemble of atoms 1.* Note that ρ_1 is a mixture.

By exploiting the physically plausible method given above for the empirical preparation of such an ensemble, an operational meaning can be attached to the process of resolution into pure subensembles. The experimental arrangement to be used is of the "coincidence" type, i.e., measurements made on an atom 2 will trigger a device accepting or rejecting the corresponding atom 1. The measurement apparatus to be used on atoms 2 is a "spin-meter" (Stern-Gerlach magnet and detection screen) which can be set to measure any spin component $\vec{\sigma}_2 \cdot \hat{n}$ by adjusting the apparatus orientation $\hat{n}$. The complete setup is schematized below:



The method of selecting subensembles involves the following steps: (a) A measurement of the observable $\vec{\sigma}_2 \cdot \hat{n}$ is performed on atom 2; this consists of its passage through the Stern-Gerlach magnet and detection in either the (+) or (-) region. (b) If detection occurs in the (-) region, the coincidence device triggers a mechanism which "opens a gate", thereby accepting the corresponding atom 1 as an element of the subensemble being selected.

Of considerable interest is the fact that an atom 1 which is accepted undergoes absolutely no physical interaction with the apparatus, the same

---

*This ρ_1 was not derived by a method which always works. The general procedure for obtaining ρ_1 from any ρ describing an ensemble of composite systems 1 and 2 is to trace over $\mathcal{H}_2$, i.e., $\rho_1 = \mathrm{Tr}_2 \rho$. To prove this, it suffices to show that

$$\mathrm{Tr}(\rho A_1 \otimes 1) = \mathrm{Tr}_1[(\mathrm{Tr}_2\rho)A_1].$$ (Cf. sec. 8.)

property attributed to the "re-arrangement" process of the preceding section. Moreover, it is easy to see that a given atom 1 would have been accepted for more than one setting $\hat{n}$ of the spin-meter.

Consider the structure of the vector $\psi = \frac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 - \beta_1 \otimes \alpha_2)$; it can be shown from P1Q and P2Q that to the results $(1,1),(1,-1),(-1,1),(-1,-1)$ of measurements of σ_{z1} and σ_{z2} performed on the ensemble of atom pairs, ψ gives the respective probabilities $0, \frac{1}{2}, \frac{1}{2}, 0$. Thus, if only the σ_{z2} -result is given, the potential σ_{z1} -result is known nevertheless. Since the σ_{z1}- and σ_{z2} -measurements obviously do not physically interfere with each other, if a σ_{z2}-measurement yields (-1) thereby triggering the acceptance of the associated atom 1, it may be concluded that a σ_{z1} -measurement would have yielded (+1). Hence, the (sub)ensemble selected in this manner is of such a nature that σ_{z1}-measurements would yield (+1) with relative frequency (probability) unity. From P1Q and P2Q it may be proved that the unique density operator which assigns probability one to σ_{z1}— measurement result (+1) is P_{α_1} *; hence the selected subensemble has density operator P_{α_1} . A glance at the structure of ρ_1 indicates that the subensemble selected in this way consists of $\frac{1}{2}$ of the original atoms 1.

Because the basic form of ρ_1 is just $\rho_1 = \frac{1}{2}1$, the ensemble of atoms 1 may be resolved into σ_{x1} -eigenvectors as follows: $\rho_1 = \frac{1}{2}P_{\delta_1} + \frac{1}{2}P_{\gamma_1}$. This means that a selection entirely different from that just described must be possible. The foregoing selection procedure involved σ_{z2} -measurements, i.e., the orientation of the Stern-Gerlach magnet was presumed set at $\hat{n} = \hat{k}$. Suppose now that $\hat{n} = \hat{\imath}$ so that the "spin-meter" now measures σ_{x2} . If there is a correlation between σ_{x1}- and σ_{x2} — measurements analogous to that between σ_{z1}- and σ_{z2}-measurements, then the same

*This is just the well known theorem that an A-eigenstate implies certainty that an A -measurement will yield the corresponding A-eigenvalue.

method used for selecting the P_{α_1} -subensemble may be adapted to the selection of a P_{δ_1}- or P_{γ_1} -subensemble. To see that such a correlation between $\sigma_{\chi 1}$ and $\sigma_{\chi 2}$ does in fact exist, ψ must be expressed in terms of δ and γ. Using the expansion of spinors α, β in terms of δ and γ, i.e., $\alpha = \frac{1}{\sqrt{2}}(\delta + \gamma)$, $\beta = \frac{1}{\sqrt{2}}(\delta - \gamma)$, we obtain

$$\psi = \frac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 - \beta_1 \otimes \alpha_2) = \frac{1}{\sqrt{2}}\left[\frac{1}{2}(\delta_1 + \gamma_1) \otimes (\delta_2 - \gamma_2) - (\delta_1 - \gamma_1) \otimes (\delta_2 + \gamma_2)\right]$$
$$= \frac{1}{\sqrt{2}}(\gamma_1 \otimes \delta_2 - \delta_1 \otimes \gamma_2).$$

Hence by setting $\hat{n} = \hat{x}$ and attaching the coincidence device to a suitably placed detector, selection of either a P_{δ_1}- or P_{γ_1} -subensemble may be accomplished.

Now, there would obviously be atoms 1 in a P_{δ_1}- or P_{γ_1} -ensemble which would have been in the P_{α_1} -ensemble had that selection been made. But these atoms 1 are in each case <u>physically independent</u> of the devices which select them. If it were meaningful to speak in a normal way of the physical "state" of a single quantum system, an atom 1 certainly could not be in a different "state" merely because a different operation was performed on a physically independent atom 2. Thus the "state" vectors of quantum theory should be ascribed no deeper significance than their role as representatives of a certain class of ensembles. Quantum systems are just never "in physical states" in the traditional sense of the phrase. We shall discuss this conclusion further in section 8.

Finally, the contradictory state assignments employed in the foregoing arguments are essentially equivalent to those made famous by the celebrated paradox of Einstein, Podolsky, and Rosen.[8] However, the present development differs from that paradox both in purpose and manner of derivation. Einstein was concerned with the quantum theory relative to a priori notions about physical reality. This interesting realm is not being explored here. Moreover, the contradictions central to the Einstein-

Podolsky-Rosen argument emerged with the help of the postulate of wave packet reduction, which was not used above. It may be that the above derivations using ensembles also lend themselves to philosophic analysis of the Einstein-Podolsky-Rosen type, but that is far afield of the present study, which is concerned with understanding quantum states rather than questioning on metaphysical grounds the completeness of the description of nature they provide.

6. The Indivisibility of Quantum Systems

"A new epoch in physical science was inaugurated...by Planck's discovery of the elementary quantum of action, which revealed a feature of wholeness inherent in atomic processes,..."[9]
--Niels Bohr

These words of the pioneer quantum philosopher from Copenhagen concisely express, by his own italics, the germinal idea from which the famous thesis of complementarity was born. Intrinsic wholeness--the indivisibility of quantum systems--is the essential ingredient of Bohr's philosophy. For almost 35 years he repeated the elements of his doctrine, at least part of which has come to be called the "orthodox" interpretation of quantum theory. However, Bohr was always content to philosophize in an almost wholly qualitative vein; even his illustrations from physics itself largely avoided the mathematics of quantum theory. Thus, in particular, the notion that quantum systems exhibit a peculiar indivisibility was always pleaded by deft application of uncertainty relations to primitive gedankenexperiments. Moreover, Bohr's depiction of quantum interaction depended strongly on an intuitive understanding of the behavior of "quanta" from the semi-classical perspective of what is often called the "Old Quantum Theory". It seems desirable therefore to provide the physical aspect of this indivisibility with a mathematical meaning in abstract quantum dynamics. First, however, in order to establish the origin of Bohr's idea of "wholeness", we review briefly some relevant points on the

philosophic side of the indivisibility concept.

As we have seen, the laws of quantum theory do not and cannot refer to anything like classical states; quantum physics is in fact characterized by an extreme nonpicturability. Put simply, quantal laws govern the statistics of measurement results, and that is all. Obviously, such a theory is not immediately reconcilable with classic aspirations of natural philosophers for an exhaustive understanding of nature with "Cartesian clarity". Prominent among the dissenters was Einstein, who regarded the irreducible probabilities of quantum theory as an intolerable weakness, a glaring sign of incompleteness. Bohr combated this feeling for decades, often in specific encounters[10] with Einstein; again and again he argued that quantum theory, understood in terms of his complementarist philosophy, is indeed exhaustive.

Bohr's elaborations of that claim typically depart from the simple philosophic observation that science is concerned with intersubjective data. Thus, however bizarre the laws of microphysics may seem, experimental contrivances must themselves be described ultimately in some communicable manner. In actual practice, classical physics provides this language in which laboratory information is phrased in objective, unambiguous, communicable facts. It is a moot point whether this use of classical constructs is necessary or conventional; Bohr seemed to favor the first alternative. In any case, the requirement of intersubjectivity is in itself not peculiar to quantum theory.

What Bohr saw as an essentially quantum feature was rather "the introduction of a fundamental distinction between the measuring apparatus and the objects under investigation",[11] a property which he regarded as a consequence of describing the apparatus in a language applicable to ordinary perceptions. As for the objects of study, quantum theory, the language

suited to their description, defies visualization, as remarked earlier. Hence, says Bohr, the interaction between apparatus and object in quantum physics is an "inseparable part of the phenomenon. Accordingly, the unambiguous account of proper quantum phenomena must, in principle, include a description of all relevant features of the experimental arrangement."[12] No longer does physical theory either permit the neglect of or offer the means to compensate for interactions with the objects; yet all knowledge of quantum objects is obtained through interactions. From this concept of indivisibility, Bohr created the principle of complementarity, according to which the totality of results of different kinds of measurements on a quantum object exhausts all conceivable knowledge about such an object, even though these results cannot be combined to form a consistent picture of that object. Further elucidation of this principle would take us too far afield, since our present subject is not complementarity itself but the related conceptual indivisibility of interacting quantum systems.

The foregoing synopsis of Bohr's idea of the "wholeness" of quantum phenomena was effectively a paraphrasing of his views with no intentional distortion. Surveying his argument, one wonders whether Bohr has outlined a description or a derivation. Indeed a first impression might suggest the latter, as though the general requirement of communicability of data in conjunction with the quantal property of nonpicturability, i.e., the failure of classical microphysics, could imply a conceptual indivisibility of interacting systems. If such an implication were truly intended, then those critics of the Copenhagen interpretation who think of Bohr and Bishop Berkeley as two of a kind are correct, for such a "derivation" of indivisibility would indeed reflect the idealist dogma, esse est percipi. There is, however, a more favorable appraisal of Bohr's conception of wholeness in quantum theory. It is the understanding of his indivisibility

notion as a philosophic description of the nature of quantum interactions, and not as a derivation of their nature from vague generalities.

To defend this proposal, let us consider the aforementioned gedanken-experiments which are typically used to enrich Copenhagen arguments by providing physical examples. Most famous of these thought experiments is Heisenberg's scheme for measuring electron position with a γ-ray microscope,[13] a token discussion of which is given in most elementary quantum mechanics textbooks. The electron position is to be observed microscopically under minimal illumination, i.e., by providing just one photon for the electron to scatter into the objective lens of the microscope. Physical optics requires short wave length illumination if a decent image of the electron is desired; but short wave length means high frequency, hence a highly energetic photon. Upon collision, such a photon would of course transfer considerable momentum to the electron; thus it is said that the position measurement affects the electron momentum in an unpredictable and uncontrollable manner, limited only by the uncertainty relation $\Delta x \Delta p \sim \left(\frac{\lambda}{\sin\theta}\right)\left(\frac{h \sin\theta}{\lambda}\right) = h$, where Δx signifies the width of the image, Δp the unknowable momentum transfer.*

Now, it is precisely this type of demonstration that Bohr takes as physical counterpart to his philosophical argument summarized above. The unpredictable effect of a position measurement on momentum suggests that the very concept of electron momentum should not even be contemplated for an electron interacting with a position-measuring device. Electron plus apparatus constitute a conceptual whole; to think of the electron independently is to divide the indivisible. So goes the complementarity argument.

Once again we ask whether this demonstration purports to derive or to

---

* θ is the angle with vertex at the electron which the scattered photon's path makes with the perpendicular from the electron to the lens.

describe. More quantitatively, does the microscope experiment explain on general grounds "why" $\Delta x \Delta p \sim h$ or is it only a picturesque illustration of the uncertainty relations for which quantum theory is assumed at the outset? The key to this question is the term unpredictable, which supposedly characterized the momentum transferred to the recoiling electron. Why is this quantity unpredictable, hence uncontrollable? Indeed Δp is calculable from the expression $\Delta p \sim \frac{h \sin\theta}{\lambda}$. The answer is simply that the quantum-theoretical approach has already been assumed; a quantal analysis of the pertinent collision problem could not predict θ. Therefore a logical textbook of quantum mechanics would not place the Copenhagen gedankenexperiments in their traditional first chapter home, but include them rather under applications of the theory (or perhaps in an appendix on the history of quantum mechanics.) Fock[14] is correct in pointing out that the mysterious term uncontrollable which is common in Copenhagen writings means nothing except that classical physics is inapplicable.

Thus it seems to us that what is really conveyed by Bohr's conception of indivisibility is a property of interaction in quantum theory, a property which can be clarified by proper identification of certain more abstract mathematical features of the theory and thereby comprehended in a manner more precise than crude inductions from gedankenexperiments. The following sections undertake such a program. First, classical statistical dynamics of interacting systems will be formulated in a manner suitable for comparison with its quantum analogue, which will be developed subsequently. Finally, from the contrast between the classical and quantal cases, there will emerge a mathematical meaning for quantum indivisibility.

7. Statistical Dynamics of Interacting Classical Systems

The basic structure of classical statistical mechanics was reviewed in sections 2 and 3. There the delta function density-of-phase,

$\rho = \delta(q-q_0)\delta(p-p_0)$, was identified as describing the pure ensemble of that theory. In the following analysis of interaction, we consider ensembles of composite systems, each consisting of two interacting subsystems. The dynamical evolution from two types of initial conditions will be examined: (1) both systems initially in pure ensembles, (2) one system initially pure, the other mixed. The fundamental problem here posed is to determine the character (pure or mixed) of the ensembles to which the two systems belong after interacting for a time. Mathematically, interaction means that the Hamiltonian function $H(q_1, p_1, q_2, p_2)$ is not additive, i.e., there do not exist $V_1(q_1, p_1)$, $V_2(q_2, p_2)$ such that $H = V_1 + V_2$.

(1) Ensembles of systems S_1 and S_2 both initially pure:

$$S_1: \rho_1(q_1, p_1; t=0) = \delta(q_1 - q_{10})\,\delta(p_1 - p_{10}).$$

$$S_2: \rho_2(q_2, p_2; t=0) = \delta(q_2 - q_{20})\,\delta(p_2 - p_{20}).$$

$$S_1 + S_2: \rho(t=0) = \rho_1(t=0)\,\rho_2(t=0).$$

It is perhaps intuitively obvious to physicists accustomed to the normal development of statistical mechanics that, under the transformation generated by the Liouville equation,

$$\rho(t=0) \longrightarrow \rho(t) = \delta(q_1 - Q_1(t))\,\delta(p_1 - P_1(t))\,\delta(q_2 - Q_2(t))\,\delta(p_2 - P_2(t)),$$

where $Q_1(t), P_1(t), Q_2(t), P_2(t)$ are the values of the canonical variables which evolve from their initial counterparts via Hamilton's equations.

Nevertheless, to facilitate comparison of classical and quantal dynamics, it is appropriate here to continue our "retrograde" development of statistical mechanics by deriving Hamilton's equations from Liouville's equation plus the concept of pure ensemble. Incidentally, although this will be done for an ensemble of pairs of interacting systems, the same method of proof is immediately adaptable to establish the simpler theorem mentioned in section 3, viz., that an initially pure ensemble of single systems evolves into a pure ensemble via the Liouville equation.

We shall now demonstrate, by direct substitution, that

$$\rho(t) = \delta(q_1 - Q_1(t))\,\delta(p_1 - P_1(t))\,\delta(q_2 - Q_2(t))\,\delta(p_2 - P_2(t)) \equiv \delta_1\delta_2\delta_3\delta_4$$

where $Q_1(0) = q_{10}$, $P_1(0) = p_{10}$, $Q_2(0) = q_{20}$, $P_2(0) = p_{20}$;

and $\dot{Q}_k = \frac{\partial H}{\partial P_k}$, $\dot{P}_k = -\frac{\partial H}{\partial Q_k}$ (Hamilton's equations), is the solution to Liouville's equation, $\frac{\partial \rho}{\partial t} = \sum\left(\frac{\partial H}{\partial q}\frac{\partial \rho}{\partial p} - \frac{\partial \rho}{\partial q}\frac{\partial H}{\partial p}\right) \equiv \{H, \rho\}$,

subject to the initial condition $\rho(t=0) = \delta(q_1 - q_{10})\,\delta(p_1 - p_{10})\,\delta(q_2 - q_{20})\,\delta(p_2 - p_{20})$.

Because confusion easily arises over the arguments of H, we establish these definitions: $H = H(q_1, p_1, q_2, p_2)$,

$$H_{q_1} \equiv \frac{\partial H}{\partial q_1},\quad H_{p_1} \equiv \frac{\partial H}{\partial p_1},\quad H_{q_2} \equiv \frac{\partial H}{\partial q_2},\quad H_{p_2} \equiv \frac{\partial H}{\partial p_2}.$$

Hamilton's equations may then be concisely expressed as

$$\dot{Q}_k = H_{p_k}(Q_1, P_1, Q_2, P_2),\quad \dot{P}_k = -H_{q_k}(Q_1, P_1, Q_2, P_2).$$

Now,

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= \frac{\partial \delta(q_1 - Q_1)}{\partial(q_1 - Q_1)}\frac{\partial(q_1 - Q_1)}{\partial t}\,\delta(p_1 - P_1)\delta(q_2 - Q_2)\delta(p_2 - P_2) \\ &\quad + \delta_1\frac{\partial \delta(p_1 - P_1)}{\partial t}\frac{\partial \delta(p_1 - P_1)}{\partial(p_1 - P_1)}\delta_3\delta_4 + \cdots \\ &= \frac{\partial \delta(q_1 - Q_1)}{\partial q_1}(-\dot{Q}_1)\delta_2\delta_3\delta_4 + \delta_1(-\dot{P}_1)\frac{\partial \delta(p_1 - P_1)}{\partial p_1}\delta_3\delta_4 + \cdots \\ &= \frac{\partial \delta_1}{\partial q_1}\left[-H_{p_1}(Q_1, P_1, Q_2, P_2)\right]\delta_2\delta_3\delta_4 + \delta_1 H_{q_1}(Q_1, P_1, Q_2, P_2)\frac{\partial \delta(p_1 - P_1)}{\partial p_1}\delta_3\delta_4 \\ &\quad + \delta_1\delta_2\frac{\partial \delta_3}{\partial q_2}(-H_{p_2})\delta_4 + \delta_1\delta_2\delta_3 H_{q_2}\frac{\partial \delta_4}{\partial p_2};\end{aligned}$$

and

$$\begin{aligned}\{H, \rho\} &= \frac{\partial \delta_1}{\partial q_1}\left[-H_{p_1}(q_1, p_1, q_2, p_2)\right]\delta_2\delta_3\delta_4 \\ &\quad + \delta_1 H_{q_1}(q_1, p_1, q_2, p_2)\frac{\partial \delta_2}{\partial p_2}\delta_3\delta_4 \\ &\quad + \delta_1\delta_2\frac{\partial \delta_3}{\partial q_2}(-H_{p_2})\delta_4 + \delta_1\delta_2\delta_3 H_{q_2}\frac{\partial \delta_4}{\partial p_2}.\end{aligned}$$

Since these expressions for $\frac{\partial \rho}{\partial t}$ and $\{H, \rho\}$ involve a generalized function (δ), their equality must be established in the sense of distribution theory. For this purpose, let $\psi_1(q_1), \phi_1(p_1), \psi_2(q_2), \phi_2(p_2)$ be testing functions. Then,

$$\int dq_1\,dp_1\,dq_2\,dp_2\,\frac{\partial \rho}{\partial t}\psi_1\phi_1\psi_2\phi_2 = -H_{p_1}\int \frac{\partial \delta_1}{\partial q_1}\psi_1\,dq_1\,\psi_2(Q_2)\phi_1(P_1)\phi_2(P_2) + \cdots$$

$$= H_{p_1}(Q_1,P_1,Q_2,P_2)\,\psi_1'(Q_1)\,\phi_1(P_1)\,\psi_2(Q_2)\,\phi_2(P_2)$$
$$- H_{q_1}(Q_1,P_1,Q_2,P_2)\,\psi_1(Q_1)\,\phi_1'(P_1)\,\psi_2(Q_2)\,\phi_2(P_2)$$
$$+ H_{p_2}\psi_1\phi_1\psi_2'\phi_2 - H_{q_2}\psi_1\phi_1\psi_2\phi_2',$$

where primes denote derivatives; and

$$\int dq_1\,dp_1\,dq_2\,dp_2\,\{H,\rho\}\psi_1\phi_1\psi_2\phi_2$$
$$= \phi_1(P_1)\psi_2(Q_2)\phi_2(P_2)\int dq_1 \frac{\partial\delta_1}{\partial q_1}[-H_{p_1}(q_1,P_1,Q_2,P_2)]\psi_1(q_1)$$
$$+ \psi_1(Q_1)\psi_2(Q_2)\phi_2(P_2)\int dp_1 \frac{\partial\delta_2}{\partial p_1} H_{q_1}(Q_1,p_1,Q_2,P_2)\phi_1(p_1)$$
$$+ \psi_1(Q_1)\phi_1(P_1)\phi_2(P_2)\int dq_2 \frac{\partial\delta_3}{\partial q_2}[-H_{p_2}(Q_1,P_1,q_2,P_2]\psi_2(q_2)$$
$$+ \psi_1(Q_1)\phi_1(P_1)\psi_2(Q_2)\int dp_2 \frac{\partial\delta_4}{\partial p_2} H_{q_2}(Q_1,P_1,Q_2,p_2)\phi_2(p_2)$$
$$= \phi_1\psi_2\phi_2\int dq_1\,\delta_1\frac{\partial}{\partial q_1}H_{p_1}(q_1,P_1,Q_2,P_2)\,\psi_1$$
$$- \psi_1\psi_2\phi_2\int dp_1\,\delta_2\frac{\partial}{\partial p_1}H_{q_1}(Q_1,p_1,Q_2,P_2)\,\phi_1 + \dots$$
$$= \phi_1\psi_2\phi_2 H_{p_1q_1}(Q_1,P_1,Q_2,P_2)\psi_1(Q_1) + \phi_1\psi_2\phi_2 H_{p_1}(Q_1,P_1,Q_2,P_2)\psi_1'(Q_1)$$
$$- \psi_1\psi_2\phi_2 H_{q_1p_1}(Q_1,P_1,Q_2,P_2)\phi_1(P_1) - \psi_1\psi_2\phi_2 H_{q_1}(Q_1,P_1,Q_2,P_2)\phi_1'(P_1) + \dots$$
$$= \phi_1\psi_2\phi_2 H_{p_1}(Q_1,P_1,Q_2,P_2)\psi_1' - \psi_1\psi_2\phi_2 H_{q_1}(Q_1,P_1,Q_2,P_2)\phi_1'$$
$$+ \psi_1\phi_1\phi_2 H_{p_2}\psi_2' - \psi_1\phi_1\psi_2 H_{q_2}\phi_2',$$

which is the same expression obtained above for $\int dq_1\,dp_1\,dq_2\,dp_2\,\frac{\partial\rho}{\partial t}\psi_1\phi_1\psi_2\phi_2$.

Hence $\frac{\partial\rho}{\partial t} = \{H,\rho\}$ for the ρ defined above.

Density functions ρ_1 and ρ_2 for the S_1- and S_2-ensembles are now determined by integration:

$$\rho_1(q_1,p_1;t) = \delta[q_1 - Q_1(t)]\,\delta[p_1 - P_1(t)],$$
$$\rho_2(q_2,p_2;t) = \delta[q_2 - Q_2(t)]\,\delta[p_2 - P_2(t)].$$

Both represent pure ensembles; therefore two initially pure ensembles remain pure regardless of mutual interaction, as expected.

(2) S_1-ensemble initially pure, S_2-ensemble initially mixed.

$$S_1:\ \rho_1(q_1,p_1;t=0) = \delta(q_1 - q_{10})\,\delta(p_1 - p_{10})$$
$$S_2:\ \rho_2(q_2,p_2;t=0) = \int w(a,b)\,\delta(q_2 - a)\,\delta(p_2 - b)\,da\,db,$$

where $w(a,b) = \rho_2(a,b;t=0)$ is the initial probability density

that an S_2 is in classical state $(q_2, p_2) = (a, b)$.

$$S_1 + S_2 : \rho(t=0) = \delta(q_1 - q_{10})\,\delta(p_1 - p_{10}) \int da\,db\, w(a,b)\,\delta(q_2 - a)\,\delta(p_2 - b)$$
$$= \int da\,db\, w(a,b)\,\delta(q_1 - q_{10})\,\delta(p_1 - p_{10})\,\delta(q_2 - a)\,\delta(p_2 - b).$$

which is just a superposition of pure subensembles.

Using the linearity of Liouville's equation, we immediately obtain $\rho(t)$ for the present case by superposition. From case (1) we have

$$\delta(q_1 - q_{10})\,\delta(p_1 - p_{10})\,\delta(q_2 - a)\,\delta(p_2 - b) \to \delta(q_1 - Q_1)\,\delta(p_1 - P_1)\,\delta(q_2 - Q_2)\,\delta(p_2 - P_2).$$

Now, each of the functions, Q_1, P_1, Q_2, P_2, depends in general on the initial parameters q_{10}, p_{10}, a, b. Therefore, none of the δ's can be removed from the integral over a, b which superposes these pure densities. Hence $\rho(t) = \int da\,db\, w(a,b)\,\delta(q_1 - Q_1)\,\delta(p_1 - P_1)\,\delta(q_2 - Q_2)\,\delta(p_2 - P_2)$, which, as is evident from its structure, represents a mixture. Moreover, both S_1- and S_2-ensembles are now mixtures: $\rho_1(t) = \int da\,db\, w(a,b)\,\delta(q_1 - Q_1)\,\delta(p_1 - P_1), \ldots$

Of special interest, the initially pure S_1-ensemble has been converted to a mixture by interaction with the initially mixed S_2-ensemble. This property of interaction, to be discussed later, will be called classical entanglement.

8. Causal Evolution of Interacting Quantum Systems

We next consider the quantum theoretical treatment of an ensemble each member of which is a composite of two interacting systems, S_1 and S_2. Mathematically, Hilbert spaces $\mathfrak{H}_1, \mathfrak{H}_2$ are associated with these systems and $\mathfrak{H}_1 \otimes \mathfrak{H}_2$ is the total Hilbert space for $S_1 + S_2$. The statement that S_1 and S_2 are interacting here means that there do not exist operators $V_1 \otimes 1$ and $1 \otimes V_2$ such that H, the total Hamiltonian operator for $S_1 + S_2$ may be written as $H = V_1 \otimes 1 + 1 \otimes V_2$.

From P3Q, the basic dynamical law of quantum mechanics is $i\hbar \frac{\partial \rho}{\partial t} = [H, \rho]$. The temporal transformation induced by this differential

equation may always be succinctly expressed by a unitary operator $T(t)$ such that $\rho(t)=T(t)\rho(0)T^{+}(t)$. Formally, $T(t)=\exp\left(-\frac{i}{\hbar}\int H dt\right)$ or when H is independent of t, $T(t)=\exp\left(-\frac{it}{\hbar}H\right)$. To prove this, it suffices to show that $\rho(t)=T(t)\rho(0)T^{+}(t)$ satisfies $i\hbar\frac{\partial\rho}{\partial t}=[H,\rho]$:

$$
\begin{aligned}
i\hbar\frac{\partial\rho}{\partial t} &= i\hbar\left[\frac{\partial T(t)}{\partial t}\rho(0)T^{+}(t)+T(t)\rho(0)\frac{\partial T^{+}(t)}{\partial t}\right] \\
&= i\hbar\left[-\frac{iH}{\hbar}T(t)\rho(0)T^{+}(t)+T(t)\rho(0)T^{+}(t)\left(\frac{iH}{\hbar}\right)\right] \\
&= H\rho(t)-\rho(t)H=[H,\rho(t)].
\end{aligned}
$$

In terms of the evolution operator $T(t)$, the statement that S_1 and S_2 are interacting means that there do <u>not</u> exist $T_1\otimes 1, 1\otimes T_2$ such that $T=T_1\otimes T_2$.

For the present problem, ρ is of course defined on $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$. To obtain ρ_1, ρ_2 , the density operators for subsystems S_1, S_2, a "partial" trace operation is required. For example, $\rho_1=Tr_2\rho$, which simply indicates a sum over those matrix elements of ρ whose two indices corresponding to $\mathcal{H}_2$ are identical. Let $\{\phi_n\}, \{\chi_m\}$ be complete orthonormal sets in $\mathcal{H}_1, \mathcal{H}_2$, respectively; then the set $\{\phi_n\otimes\chi_m\}$ spans $\mathcal{H}_1\otimes\mathcal{H}_2$. A typical matrix element for ρ is then $\rho_{kelmn}=\langle\phi_k\otimes\chi_e|\rho|\phi_m\otimes\chi_n\rangle$, and $(\rho_1)_{em}=\sum_n \rho_{enmn}$; or $\rho_1=\sum_m\langle\chi_m|\rho|\chi_m\rangle=Tr_2\rho$, symbolically.

To prove that $\rho_1=Tr_2\rho$ is the density operator for an ensemble of subsystems S_1 , first note that all Hermitean operators representing observables belonging exclusively to S_1 have the form $A_1\otimes 1$. Accordingly, the mean value functional for S_1 has the following form:

$$
\begin{aligned}
m(A_1\otimes 1) &= Tr(\rho A_1\otimes 1)=\sum_{nm}\langle\phi_n\otimes\chi_m|\rho A_1\otimes 1|\phi_n\otimes\chi_m\rangle \\
&= \sum_{nm}\langle\phi_n\otimes\chi_m|\rho|A_1\phi_n\otimes\chi_m\rangle \\
&= \sum_n\langle\phi_n|\sum_m\langle\chi_m|\rho|\chi_m\rangle A_1|\phi_n\rangle \\
&= Tr_1[(Tr_2\rho)A_1]=Tr_1(\rho_1 A_1).
\end{aligned}
$$

Hence, $\rho_1=Tr_2\rho$ is the density operator for the S_1-ensemble.

A detailed exposition of the theory of composite systems may be found

in von Neumann's treatise.[15] There the following theorems, the first two of which will be used below, are proved:

(a) If ρ_1 and ρ_2 are both pure, $\rho = \rho_1 \otimes \rho_2$ uniquely.

(b) If either ρ_1 or ρ_2 is pure, $\rho = \rho_1 \otimes \rho_2$ uniquely.

(c) If ρ_1 and ρ_2 are both mixed, ρ is not uniquely determined.

Now, working in parallel with the preceding section we shall determine the character (pure or mixed) of ensembles of interacting quantum systems $\underline{S}_1$ and $\underline{S}_2$.

(1) Ensembles of systems $\underline{S}_1$ and $\underline{S}_2$ both initially pure.

$$\underline{S}_1 : \rho_1(t=0) = P_{\phi_0}$$
$$\underline{S}_2 : \rho_2(t=0) = P_{\chi_0}$$
$$\underline{S}_1 + \underline{S}_2 : \rho(t=0) = P_{\phi_0} \otimes P_{\chi_0} = P_{\phi_0 \otimes \chi_0}.$$

The evolution operator $T(t)$ transforms $\rho(t=0)$ into

$$\rho(t) = T(t) P_{\phi_0 \otimes \chi_0} T^{\dagger}(t) = |T(\phi_0 \otimes \chi_0)\rangle\langle T(\phi_0 \otimes \chi_0)| = P_{T(\phi_0 \otimes \chi_0)}.$$

Thus, as in the classical case, the overall homogeneity of the ensemble is not altered by causal evolution. The density operators for the $\underline{S}_1$- and $\underline{S}_2$-ensembles are now determined by the partial trace equations. For this calculation, consider the expansion

$$T(\phi_0 \otimes \chi_0) = \sum_{nm} C_{nm} \phi_n \otimes \chi_m \equiv \psi.$$

Now,
$$\rho_1(t) = \mathrm{Tr}_2 P_{\psi}$$
$$= \sum_{kr} \langle \chi_{kr} | \sum_{nm} C_{nm} \phi_n \otimes \chi_m \rangle \langle \sum_{ls} C_{sl} \phi_s \otimes \chi_l | \chi_{kr} \rangle$$
$$= \sum_{knmls} \delta_{lk} \delta_{rm} C^{*}_{sl} C_{nm} |\phi_n\rangle\langle\phi_s|$$
$$= \sum_{nrs} C^{*}_{sr} C_{nr} |\phi_n\rangle\langle\phi_s| \quad , \text{ a mixture}$$

(since this is not in general a projection operator).

If $\underline{S}_1$ and $\underline{S}_2$ did <u>not</u> interact, i.e., $H = V_1 \otimes 1 + 1 \otimes V_2$, then $\exp[-\frac{it}{\hbar} H](\phi_0 \otimes \chi_0) = \exp(-\frac{it}{\hbar} V_1)\phi_0 \exp(-\frac{it}{\hbar} V_2)\chi_0$, since $\exp(A+B) = \exp A \exp B$ if $[A,B]=0$. Hence C_{nm} would have the product form $C_n C_m$, and then

$$\rho_1(t) = \sum_{nks} C_s^* C_k^* C_n C_{ok} |\phi_n\rangle\langle\phi_s|$$
$$= |\sum_n C_n \phi_n\rangle\langle\sum_s C_s \phi_s| \left(\sum_k |C_{ok}|^2\right) = P_\phi,$$

where $\phi = \sum_n C_n \phi_n = \exp\left(-\frac{it}{\hbar} V_1\right)\phi_0$.

Similar results hold for $\rho_2(t)$.

To summarize: unlike the classical case, two initially pure quantum ensembles are generally converted to mixtures during interaction; however, in the absence of interaction, both remain homogeneous. In the following section, we shall refer to this peculiarly quantum mixing as <u>quantum entanglement</u>.

(2) $\underline{S}_1$ – ensemble initially pure, $\underline{S}_2$ – ensemble initially mixed.

$$\underline{S}_1 : \rho_1(t=0) = P_{\phi_0}$$
$$\underline{S}_2 : \rho_2(t=0) = \sum_k w_k P_{\eta_k}, \quad \sum_k w_k = 1.$$
$$\underline{S}_1 + \underline{S}_2 : \rho(t=0) = P_{\phi_0} \otimes \sum_k w_k P_{\eta_k}.$$

As in the classical case (2), linearity of the dynamical law enables solution by superposition. Thus, from $T(t)\left(P_{\phi_0} \otimes P_{\eta_k}\right)T^+(t) = \rho_k(t) = P_{\psi_k}$, where $\psi_k = \sum_{nm} C_{nm}^{(k)} \phi_n \otimes \chi_m$, it follows that $\rho(t=0) \longrightarrow \rho(t) = \sum_k w_k P_{\psi_k}$; i.e., ρ, initially a mixture, in general remains inhomogeneous.

Of great interest here is the effect of interaction on the initially pure $\underline{S}_1$-ensemble. Taking the appropriate trace, we find, using the result from case (1),

$$\rho_1(t) = Tr_2 \rho(t) = \sum_k w_k Tr_2 P_{\psi_k}$$
$$= \sum_k w_k \left(\sum_{rsn} C_{sr}^{(k)*} C_{nr}^{(k)} |\phi_n\rangle\langle\phi_s|\right),$$

which in general represents a mixture.

To clarify the nature of this latter transformation, $P_{\phi_0} \longrightarrow \sum_k w_k \left(\sum_{rsn} C_{sr}^{(k)*} C_{nr}^{(k)} |\phi_n\rangle\langle\phi_s|\right)$, let us suppose quantum entanglement did not exist. Then case (1) would have yielded

$$P_{\phi_0} \otimes P_{\eta_k} \longrightarrow P_{\phi(\phi_0, \eta_k)} \otimes P_{\eta(\phi_0, \eta_k)}$$; superposing these, we obtain

$$\rho(t) = \sum_k w_k P_{\phi(\phi_0, \eta_k)} \otimes P_{\eta(\phi_0, \eta_k)}.$$

Thus $\rho_1(t) = \sum_k w_k \sum_n \langle \chi_n | P_\phi \otimes P_\eta | \chi_n \rangle$

$= \sum_k w_k P_{\phi(\phi_0, \eta_k)} Tr_2 P_{\eta_k} = \sum_k w_k P_{\phi(\phi_0, \eta_k)}$,

which we recognize as analogous to the effect previously called classical entanglement.

Returning now to the correct result,

$$\rho_1(t) = \sum_k w_k \left(\sum_{rsn} C_{sr}^{(k)*} C_{nr}^{(k)} |\phi_n\rangle\langle\phi_s| \right),$$

we identify the mixtures in parentheses as a consequence of quantum entanglement, while the further mixing by the sum over k corresponds to classical entanglement.

9. Classical Divisibility and Quantum Indivisibility from a Dynamical Viewpoint

In a certain sense, the foregoing deductions (sec. 7) concerning classical interactions are paradoxical. For any Hamiltonian whatsoever, even one which includes interaction terms, the homogeneity of an ensemble of systems interacting with members of a second initially pure ensemble is preserved throughout the motion; nevertheless, by arranging an interaction between the members of an initially pure ensemble and those of a mixture, the initially homogeneous ensemble in general changes into an inhomogeneous ensemble after a time, a process we called classical entanglement. Therefore we confront a strange dilemma, for in classical entanglement mere juxtaposition apparently achieves the impossible, viz., the destruction of homogeneity, which no Hamiltonian, hence no force conceivable within this theory, can accomplish. To explain this seeming discrepancy, consider a simple problem in classical physics.

Let S_1 and S_2 each be a particle of mass m, interacting via a connecting spring of stiffness k and equilibrium length ℓ. The Hamiltonian function is therefore $H = \frac{1}{2m}(p_1^2 + p_2^2) + \frac{k}{2}(q_2 - q_1 - \ell)^2$,

where p_1, p_2 denote linear momenta and q_1, q_2 are position co-ordinates for S_1, S_2. From Hamilton's equations it is easy to obtain Newton's law:

$$m\ddot{q}_1 = k(q_2 - q_1 - \ell), \quad m\ddot{q}_2 = -k(q_2 - q_1 - \ell).$$

When the latter pair of coupled differential equations are solved subject to the initial conditions $q_1(0) = 0$, $\dot{q}_1(0) = \dot{q}_2(0) = 0$, $q_2(0) = q_{20}$ the result is $q_1(t) = \frac{1}{2}(q_{20} - \ell)[1 - \cos\sqrt{2}\,\omega t]$, $\omega = \sqrt{\frac{k}{m}}$.

Now, a relevant question is whether there exists some force which, when acting on S_1 produces the same motion $q_1(t)$ as did the above interaction. Certainly. In fact the Hamiltonian

$$H_1 = \frac{p_1^2}{2m} + \frac{1}{2}(2k)\left(q_1 - \frac{q_{20} - \ell}{2}\right)^2$$

produces the desired motion. (Physically, S_1 could be at one end of a spring of stiffness 2k and equilibrium length $\frac{q_{20} - \ell}{2}$ which is rigidly mounted at its other end.)

Consider an initial S_2-ensemble mixed by being distributed over the values of q_{20}. By classical entanglement, an initially pure S_1-ensemble interacting with this mixed one will itself become inhomogeneous. The reason for this is now easily explained by the observation that the Hamiltonian H_1 <u>depends on</u> q_{20}. Thus the S_1-"ensemble", when considered by itself, is not an ordinary <u>ensemble</u> at all; for it does not consist of a collection of identical systems but rather of a distribution of <u>different kinds of systems</u>, characterized by <u>different</u> Hamiltonians $H_1(q_{20})$. (Physically, the various S_1 could be attached to springs with differing equilibrium lengths.) It is therefore not alarming that this "ensemble" behaves in a manner contradictory to the Liouville equation.

The paradox of classical entanglement was illusory; the "impossible" disruption of homogeneity in the S_1-ensemble had a theoretical explanation. Thus, the existence of classical entanglement is no reason to declare interacting systems S_1 and S_2 to be in any sense indivisible.

where p_1, p_2 denote linear momenta and q_1, q_2 are position co-ordinates for S_1, S_2. From Hamilton's equations it is easy to obtain Newton's law:

$$m\ddot{q}_1 = k(q_2 - q_1 - \ell), \quad m\ddot{q}_2 = -k(q_2 - q_1 - \ell).$$

When the latter pair of coupled differential equations are solved subject to the initial conditions $q_1(0) = 0$, $\dot{q}_1(0) = \dot{q}_2(0) = 0$, $q_2(0) = q_{20}$ the result is $q_1(t) = \frac{1}{2}(q_{20} - \ell)[1 - \cos\sqrt{2}\,\omega t]$, $\omega = \sqrt{\frac{k}{m}}$.

Now, a relevant question is whether there exists some force which, when acting on S_1 produces the same motion $q_1(t)$ as did the above interaction. Certainly. In fact the Hamiltonian

$$H_1 = \frac{p_1^2}{2m} + \frac{1}{2}(2k)\left(q_1 - \frac{q_{20} - \ell}{2}\right)^2$$

produces the desired motion. (Physically, S_1 could be at one end of a spring of stiffness 2k and equilibrium length $\frac{q_{20} - \ell}{2}$ which is rigidly mounted at its other end.)

Consider an initial S_2-ensemble mixed by being distributed over the values of q_{20}. By classical entanglement, an initially pure S_1-ensemble interacting with this mixed one will itself become inhomogeneous. The reason for this is now easily explained by the observation that the Hamiltonian H_1 <u>depends on</u> q_{20}. Thus the S_1-"ensemble", when considered by itself, is not an ordinary <u>ensemble</u> at all; for it does not consist of a collection of identical systems but rather of a distribution of <u>different kinds of systems</u>, characterized by <u>different</u> Hamiltonians $H_1(q_{20})$. (Physically, the various S_1 could be attached to springs with differing equilibrium lengths.) It is therefore not alarming that this "ensemble" behaves in a manner contradictory to the Liouville equation.

The paradox of classical entanglement was illusory; the "impossible" disruption of homogeneity in the S_1-ensemble had a theoretical explanation. Thus, the existence of classical entanglement is no reason to declare interacting systems S_1 and S_2 to be in any sense indivisible.

For quantum entanglement, the situation is quite different. Here no explanation parallel to that given for the classical case is possible. Quantal entanglement occurs under conditions of maximal homogeneity (S_1 and S_2 both initially pure); but it was the inhomogeneity of S_2 and its reflection in H_1 that made classical entanglement "divisible", hence explicable. In the quantum case there is no way to explain away the basic paradox of entanglement, viz., that interacting systems S_1 and S_2 each from initially pure ensembles develop temporally into members of mixed ensembles. Since there does not exist an evolution operator T such that $T P_{\varphi_0} T^{+} = \rho_t$ where ρ_t is not a projection operator, no physical environment conceivable within quantum theory has the same effect on the initially pure S_1-ensemble as quantum entanglement with the S_2-ensemble. Thus, in an explicit dynamical sense, interaction in quantum theory exhibits a remarkable property of "wholeness", to use Bohr's word.

If S_1 is an atomic system and S_2 the measuring device through which S_1 is studied via interaction, then it is in fact a quantum dynamical property of this interaction that it becomes impossible--even with quantum theory--to give an independent account of the temporal development of S_1. The composite system $S_1 + S_2$, quantally entangled, may therefore be regarded as dynamically "indivisible". With this mathematical interpretation, perhaps Bohr's concept of "wholeness", a fundamental attribute of quantum theory and a pillar of complementarity, obtains a more definitive meaning.

10. Summary: The State Concept in Quantum Theory

Ever since Born first provided quantum theory with its fundamental link to nature via statistics, controversy has raged over the extent to which this innovation modifies the basic classical structure of physical

science. Since the only quantal constructs which participate in a causal law relate to nature solely through probabilistic-statistical rules of correspondence, it is obvious that any "state" concept in quantum theory must refer empirically to statistical ensembles instead of individual systems. By this we mean simply that the "states" of quantum theory are related to statistical collectives of measurement results emerging from measurements upon identically prepared systems. The only sense in which such "states" might be construed as referring empirically to a single system is in the case of an ensemble consisting of one system sequentially measured and reprepared; however, this is beside the point in the present inquiry, which has sought to ascertain whether or not one may consistently regard quantum "states" as belonging to physical systems in the classical manner wherein every system is thought of as always being in some definite (possibly unknown) state. The fact that quantum states refer empirically to ensembles does not preclude the theoretical possibility of restoring a causal nexus for individual states, as was demonstrated above for the case of "retrograde" classical statistical mechanics. However, that possibility was found to hinge upon (1) the question of uniqueness in the problem of resolution of general ensembles into pure subensembles, since the pure ensemble is the appropriate construct from which to develop an individual state specification for an initially statistical theory; and (2) the question of "conservation of homogeneity" for an initially pure ensemble under causal evolution, since any meaningful individual state concept should be applicable to a system at all times, i.e., it should be possible to follow the temporal development of the state via the causal law of the theory.

The preceding sections sought the meaning of quantum states within the minimal logical framework of quantum mechanics; accordingly, the analysis centered on the ensembles which give rise to the statistical collectives

of measurement results actually studied in quantum theory. No quantum state concept was assumed a priori; instead we began only with the elementary assumption that there are systems upon which measurements of observables are performed according to established rules of correspondence, and that the statistics of the numerical results which emerge from such measurements upon an ensemble of identically prepared systems are governed by quantum theory, (P1Q, P2Q, P3Q). Against this background the logical standing of the usual theoretical quantum state concept was assayed using the above criteria (1) and (2).

The results of the analysis, with comparisons to the analogous classical situation, may be summarized as follows:

(1) Decomposition of the classical density-of-phase into representatives of pure subensembles is unique; thus single classical systems may be assigned states (q_0, p_0) unambiguously. On the other hand, resolution of quantum ensembles into pure subensembles is not unique; hence, as illustrated earlier, the assignment of a state vector ψ to a single system at a single time is an ambiguous procedure which can lead to theoretical paradoxes.

(2) A classical pure ensemble, even when interacting with another pure ensemble, remains at all times pure; it makes sense therefore to assign a temporal sequence of states $(q(t), p(t))$ to any classical system. On the contrary, although a pure ensemble of closed quantum systems does conserve its homogeneity, a quantum pure ensemble interacting with another pure ensemble becomes entangled and is converted to a mixture; thus it is in general impossible to assign a temporal sequence of state vectors $\psi(t)$ to a quantum system.

In short, quantum theory satisfies neither of the two criteria necessary to justify the supplementation of an initially statistical theory by

a state concept in the classic sense. Although classical statistical mechanics admits of an unambiguous assignment of individual states, quantum theory fails to satisfy the necessary criteria. Hence the simplest, and most natural, conclusion is that the pure "state" vector ψ of quantum theory must not be interpreted--even theoretically--as referring to the physical state of a single system at a single time. This does not imply that the basic meaning of causality is lost in quantum physics. Although the classical ideal of determinism as applied to single physical events is not valid in quantum mechanics, the behavior of quantum ensembles is predictable in the sense that future measurement statistics are determined by past measurement statistics. Thus, strictly speaking, the state concept in quantum mechanics belongs to the ensemble instead of the system. Indeed quantum systems should be regarded as never being in any physical state (except in the aforementioned statistical sense where the "state" refers to a single system because the ensemble consists of one system sequentially measured and reprepared). To use a terminology sometimes used in statistical theories, quantum mechanics may be characterized as a theory with macrostates (of ensembles) for which there are no underlying microstates (of systems).

A highly metaphorical illustration of this conclusion may be obtained by contrasting the classical and quantum descriptions of the following gedankenexperiment. Let the object of study (the physical system) be a "tossed coin" prepared for measurement by feeding it into one of several different kinds of tossing devices $\Pi_1, \Pi_2, \ldots$, each of which ejects it onto a tabletop divided into two regions. Consider the observables $\mathcal{R}$ and $\mathcal{S}$ associated with the tossed coin and defined by the following questions: $\mathcal{R}$: Did the coin land in region 1, or in region 2?; $\mathcal{S}$: Did the coin land "heads up" or "tails up"? Now, if we select one of the tossing

devices, say Π_1, then feed the coin into it, measure the observables of the tossed coin, and repeat the procedure many times, we can obtain probability distributions $W'_{\mathcal{R}}, W'_{\mathcal{S}}$ for $\mathcal{R}$- and $\mathcal{S}$-measurements upon a tossed coin prepared by method Π'. (The results will be the same regardless of whether one coin is recycled or many identical coins are fed through.) If another tossing device, say Π^2, is used, different distributions $W^2_{\mathcal{R}}, W^2_{\mathcal{S}}$ will result.

If the coin were <u>classical</u>, the measurement statistics for $\mathcal{R}, \mathcal{S}$ could be summarized by a "density function" $W^k(R,S)$, where $k = 1, 2, \ldots$ refers to the method of preparation Π_k. Since the experiments which $W^k(R,S)$ describes may be performed by recycling just one coin, in a certain sense $W^k(R,S)$ may be associated with a single coin, or more precisely, with a single coin <u>plus</u> a definite method of preparation Π_k. $W^k(R,S)$ is, so to speak, the "state" of an ensemble prepared in the manner Π_k. However, $W^k(R,S)$ cannot be regarded as the physical condition, or state, of <u>a</u> coin, for apart from the context of an ensemble prepared by repeated application of Π_k, it would be meaningless to say that every tossed coin is in some unique "state" $W^k(R,S)$. To find an acceptable state concept for individual systems, the pure ensembles must be identified and their properties analyzed in the manner described in preceding sections. For a classical coin such an analysis would show that every tossed coin may without contradiction be regarded as always having definite values (perhaps unknown) of $\mathcal{R}$ and $\mathcal{S}$, i.e., as being in either region 1 or 2 with either heads up or tails up. Some pair of $\mathcal{R}, \mathcal{S}$ values could therefore represent the <u>state</u> of any classical coin.

If, on the other hand, the coin were a <u>quantum</u> object, the measurement statistics for $\mathcal{R}, \mathcal{S}$ could be summarized by a "density operator" ρ^k. As in the classical case, there is a certain sense in which ρ^k may be associated with a single coin (<u>plus</u> Π_k); but once again it would be

meaningless to say that every tossed coin is in some unique "state" ρ^{R} unless this statement were qualified by reference to the ensemble produced by repeated application of T_R. Thus ρ^{R} represents the "state" of an ensemble prepared in the manner T_R. Now, the remarkable feature of quantum theory with which this paper has been concerned is that the latter "state of an ensemble" is the only possible quantum state concept. Although there are pure ensembles in quantum mechanics (represented by the well known state vectors, or wave functions), these pure ensembles do not lead to a consistent state concept applicable to single systems. Not only is there no classical state concept in quantum theory, i.e., the notion of a tossed coin having definite R, S values does not appear even among the pure ensembles (assuming $[R,S]\neq 0$), there is no individual state concept at all in quantum theory. It is not even permissible to regard every tossed quantum coin as having a definite, but perhaps unknown, state vector. State vectors merely correspond to a certain kind of density operator and as such represent "states" only in the statistical sense in which all density operators represent"states" of ensembles prepared in some given manner T.

To summarize, a classical tossed coin has both (1) a probabilistic "macrostate" $W^{R}(R,S)$ which depends on the nature of the experiment used to study the coin, i.e., upon T_R, and also (2) a definite state, the "microstate" expressed by a pair of values for R, S which represent the intrinsic physical condition of the coin. A quantum tossed coin, on the other hand, has only (1) a probabilistic "macrostate" ρ^{R}, which may be pure $(\rho^{R}=P_{\psi_R})$, but there is no quantum analogue whatsoever for (2).

Finally, perhaps it should be emphasized that the purpose of this analysis was not at all semantic; i.e., no attempt was being made above to provide physicists with a better jargon for their workaday activities.

Everyday statements made about individual systems are usually harmless since they are easily translatable into the proper corresponding assertions about ensembles. However, in deeper theoretical and philosophical considerations, the problems raised above concerning the meaning of quantum states do transcend mere verbal maneuvers. In particular, a satisfactory understanding of the state concept in quantum theory is pre-requisite to any rational study of general quantum measurement theory and to any thoughtful evaluation of generalized quantum ideas at the contemporary frontiers of physics.

Selected References

1. H. Margenau, The Nature of Physical Reality, McGraw-Hill (1950).

2. Richard von Mises, Mathematical Theory of Probability and Statistics, Academic Press (1964), 43-49; cf. also Margenau, op. cit., Ch. 13.

3. H. Margenau, Philosophy of Science, 4 (1937), 352-356.

4. J. L. Park, "Quantum Theoretical Concepts of Measurement" (Part II of Thesis*); "Simultaneous Measurability in Quantum Theory (Part III of Thesis*).

5. J. von Neumann, Mathematical Foundations of Quantum Mechanics, Eng. trans. by R. T. Beyer, Princeton (1955), 295-313.

6. Ibid. 321-323.

7. E. Schrödinger, Proc. Cambr. Philo. Soc., 32 (1936), 446.

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

9. N. Bohr, Essays 1958-1962 on Atomic Physics and Human Knowledge, Interscience (1963), 2.

10. Bohr, Atomic Physics and Human Knowledge, Wiley (1958), 32.

11. Bohr, Essays 1958-1962, 3.

12. Ibid., 4.

13. W. Heisenberg, The Physical Principles of the Quantum Theory, Univ. of Chicago (1930), 21.

14. V. Fock, Czechoslovakian Journal of Physics, 7 (1957), 643.

15. von Neumann, op. cit., 422-437.

---

*Because the three self-contained parts of this dissertation were written with separate publication in mind, each part treats the other parts as independent references.

II. QUANTUM THEORETICAL CONCEPTS OF MEASUREMENT

ABSTRACT

The overall purpose of Part II is to clarify the physical meaning and epistemological status of the term measurement as used in quantum theory. A general introduction to the measurement problem is followed first by a review of the essential logical structure of quantum physics, with due emphasis on the conclusions of Part I, and by interpretive discussions contrasting the quantal concepts observable and ensemble with their classical ancestors along the lines of Margenau's latency theory. Against this background various popular ideas concerning the nature of quantum measurement are critically surveyed. The analysis reveals that, in addition to internal mathematical difficulties, all the so-called quantum theories of measurement are grounded in unjustifiable, classical presuppositions. After a sequence of critiques, the remainder of Part II seeks to develop an acceptable quantal understanding of the concepts measurement and preparation.

A careful study of the quantum description of real experiments is used to motivate a proposal that two distinct quantum theoretical measurement constructs should be recognized, both of which must be distinguished from the concept of preparation. The different epistemological roles of these concepts are compared and explained. It is then concluded that the only possible type of "quantum measurement theory" is one of little metaphysical interest and that quantum measurement seems problematical only when viewed from an overly narrow classical perspective.

1. Measurement

Theoretical physics draws its remarkable potency for explanation and prediction chiefly from its fusion with mathematics, a union which confers upon physics the notable advantages inherent in the logical manipulation of concepts within a deductive framework. Such deductive machinery working in tandem with inductive experimental methods epitomizes the scientific method. It is the process called measurement through which the theoretical and empirical aspects are linked. Accordingly, that process must never be forgotten in stating the postulates for any physical theory.

Prior to the quantum era, the measurement concept was philosophically innocuous; it displayed a certain obviousness of meaning which occasioned little controversy. In fact, the postulates of classical theories made reference to it only implicitly. However, for reasons to be discussed below in connection with the related concept of observable, quantal propositions cannot suppress direct use of the term.

In spite of its great strength as a bulwark of theoretical physics, mathematization is ever haunted by the ghost of potential logical inconsistency; it is quite possible to form from a given set of physical concepts some plausible axioms which yield numerous empirically verified deductions yet lead also to logical contradictions, in which case the explanatory value of the theory is sharply diminished.

Because quantum theory involves radical departures from classical modes of thought regarding nature and natural law and, more explicitly, because its statements use old scientific terms in new combinations, the logical consistency of the quantal algorithm may be reasonably challenged. Roughly speaking, such objections fall into two categories which we may call mathematical and philosophical. The former embraces such important technical difficulties as the old problem of rigorous justification of

Dirac delta-functions and the newer perplexities of renormalization in quantum field theory. Such problems are not, however, peculiar to quantum theory; classical physics had, indeed still has, its own mathematical blemsihes, e.g., the many body problem of classical mechanics, renormalization in classical electrodynamics, and the ergodic problem in statistical mechanics. At any rate, we are not here concerned with problems of this type but focus instead on the second category of logical challenges, the philosophical ones. Generally, these may be expected to persist intact regardless of whatever progress is made toward resolving the technical mathematical dilemmas. What are here called philosophical challenges are by no means nonmathematical, however; the essential point is that they are not merely mathematical. For example, we shall work with vector spaces and associated operators, although the precise mathematical foundations of these theoretical structures as employed by quantum theorists are not yet rigorously grounded. There is a growing literature in this realm of "quantum mathematics", but we shall be concerned only with those logical features common to all such endeavors, i.e., with general postulates which reflect the essence of quantum theory independently of the choice of mathematical background.

As noted above, quantal postulates differ from classical ones in that the formerly tacit concept of measurement has emerged to take part explicitly in physical statements. It is natural, therefore, in probing the quantum framework for logical inconsistencies, to seize upon this novel feature by demanding a consistent quantal description of the process of measurement; the replies given to this logical challenge are called quantum theories of measurement. Because measurement is not a concept in isolation, the study of such theories reveals diverse ideas concerning the nature of other quantal constructs. Hence the quantum theory of measurement offers

a portal to philosophic understanding of the meaning and goals of quantum physics as a whole.

2. Mathematical Foundations of Quantum Physics

To provide a firm basis for our analysis of measurement, we first review the fundamental axioms of quantum physics. In the spirit of the foregoing remarks, our postulates attempt to capture the essential character of a quantum theory, not to enumerate every single mathematical assumption. This is standard practice in physics; we force an admittedly arbitrary distinction between unstated background postulates, which encompass much logic and mathematics, and physical postulates, which serve to delineate the rudiments of a particular branch of physics. Such a cleavage enables us to direct our philosophic inquiry more acutely to crucial physical points instead of detracting from that purpose by citing numerous minor axioms in the manner of the more tedious excursions into "quantum mathematics"[1].

We now state and discuss three postulates which underlie all forms of modern quantum theory from wave mechanics to field theory, postulates which reflect the essence of the quantum approach to natural philosophy. From these we shall extract the primitive physical terms employed, and our analysis of measurement will then revolve about those basic constructs.

P1: (Correspondence Postulate) The* linear Hermitean operators, A, B,..., on Hilbert space which have complete orthonormal sets of eigenvectors correspond to physical observables $\mathcal{A}, \mathcal{B}, \ldots$ The

---

*To accommodate superselection rules[2], principles which in one form prohibit certain Hermitean operators from representing observables, the initial the in P1 might have to be replaced by some (cf. sec. 19).

function $\mathcal{F}(A)$ corresponds to observable $\mathcal{F}(a)$ if A corresponds to a.*

The correspondence postulate is often stated in the converse form: to every observable there corresponds an operator. However, we have shown elsewhere[3] that such a formulation leads to physically untenable consequences and must be rejected. The term Hilbert space appears due to established physical usage; no stricture is intended on the application of newer mathematical constructs which may eventually provide the mathematical background for quantum theory. This exemplifies a point made at the outset: regardless of the precise mathematical schema which becomes the "Hilbert space", the physical and philosophic meaning of the correspondence postulate remains the same.

P2: (Mean Value Postulate) To every ensemble of identically prepared systems there corresponds a real linear functional of the Hermitean operators, m(A), such that if A corresponds to an observable a, the value of m(A) is the arithmetic mean of the results of a-measurements performed on the member systems of the ensemble.

P3: (Dynamical Postulate) Every type of quantum system is dynamically characterized by a linear unitary operator T (the evolution operator) in the sense that the mean value functional $m_{t_2}(A)$ at time t_2 for an ensemble of such systems which at time t_1 had mean value functional $m_{t_1}(A)$ is given by

$$m_{t_2}(A) = m_{t_1}\left[T^{\dagger}(t_2, t_1) A T(t_2, t_1)\right].$$

P2 and P3 together indicate that the concept of physical state in quantum theory is represented by the mean value functional, the only quantal

* The observable $\mathcal{F}(a)$ is measured simply by measuring a and substituting the result a into the function $\mathcal{F}$; the range value, $\mathcal{F}(a)$, is then the result of the $\mathcal{F}(a)$-measurement.

construct which relates to measurement results and obeys a causal law. Classically, this is perhaps the most objectionable feature of quantum theory, for m(A) refers empirically only to an ensemble whereas a state representation traditionally belonged to individual systems in a nonstatistical sense. However, a cursory examination of these postulates does not immediately show that the older understanding of the state concept cannot somehow be extricated from them, although this is in fact the case.[4] We shall return to this point later.

To make contact with familiar elements of the quantum formalism, we next state a few key theorems which follow[3] from P1-P3.

Th1:[5] For every mean value functional m(A) there exists an Hermitean operator ρ such that

$$m(A) = Tr(\rho A).$$

For mathematical convenience, it is fruitful to shift the emphasis from the functional m(A) to the operator ρ related to it by Th1. Thus, the statistical properties of an ensemble are embodied in ρ, which is called the density operator.

Th2: $Tr\rho = 1$.

Th3: The probability $W_a(a_k;\rho)$ that an a-measurement on a system from an ensemble with density operator ρ will yield eigenvalue a_k; $A\alpha_{kd_k} = a_k \alpha_{kd_k}$ is given by

$$W_a(a_k;\rho) = Tr(\rho P_{\mathcal{L}_k}),$$

where $P_{\mathcal{L}_k}$ is the projection operator onto the subspace $\mathcal{L}_k$ belonging to eigenvalue a_k: $P_{\mathcal{L}_k} \equiv \sum_{d_k} P_{\alpha_{kd_k}}$.

Th4: The only possible results of a-measurements are the eigenvalues of corresponding operator A.

Th5: The density operator ρ is positive semidefinite.

Of tremendous significance in the theory of measurement is the concept

of ensemble homogeneity, emphasized by von Neumann.[6] For the present, we merely define it, deferring philosophic analysis to later sections.

Defn: An ensemble is said to be pure, or homogeneous, if every rearrangement and partitioning of member systems results in subensembles physically identical to the original. An ensemble which is not pure is said to be mixed, or a mixture.

In terms of the mean value functional, if m(A) is pure, there do not exist $m_1(A), m_2(A)$ satisfying $m(A) = w_1 m_1(A) + w_2 m_2(A)$, where w_1, w_2 are the fractions of the original ensemble contained in the two subensembles; clearly, $w_1 + w_2 = 1$, $w_1 > 0$, $w_2 > 0$. In the language of density operators, we then have the following theorem:

Th6: ρ is pure if and only if $\rho = P_\psi$, where P_ψ is the projection operator onto the span of Hilbert vector ψ.

(ψ is usually called the "state" vector.)

Th7: For pure ensembles with state vector ψ, i.e., $\rho = P_\psi$, $m(A) = \frac{\langle \psi, A\psi \rangle}{\langle \psi, \psi \rangle}$, where $\langle , \rangle$ denotes scalar product.

By convention, ψ is generally normalized so that $\langle \psi, \psi \rangle = 1$, hence $m(A) = \langle \psi, A\psi \rangle$.

The theorems above comprise the basic ingredients of quantum statics, so called because all statements essentially refer to a single instant of time. Quantum causality is embodied in the temporal development of m(A) according to P3. From Th1, we have $m(A) = Tr(\rho A)$; hence, temporal changes of the functional m may be represented by transformations either of the density operator (Schrödinger picture), or of the operators representing observables (Heisenberg picture), or of both. In what follows, quantum dynamics will be cast in the Schrödinger picture.

Th8: $\rho(t_2) = T(t_2, t_1)\rho(t_1)T^\dagger(t_2, t_1)$.

Th9: If $\rho(t_1) = P_{\psi(t_1)}$, then $\rho(t_2) = P_{\psi(t_2)}$, where

$\psi(t_2) = T(t_2, t_1)\,\psi(t_1)$.*

Th7 and Th9 provide the justification for the common name "state" vector for ψ, since knowledge of ψ and T enables calculation of all measurement statistics for any instant.

Dominating the contemporary frontier of theoretical quantum physics is a construct not yet mentioned, the quantum field. Its omission from the foregoing postulates should not, however, be construed as any serious limitation upon their domain of physical relevance. Actually, quantum field theory is easily subsumed under these general quantal postulates which, it will be recalled, were intended to reflect the skeletal structure of any "quantum" theory exclusive of the details of a particular realization. Thus, for example, under the Correspondence postulate (P1), statements of the various commutation relations which serve to define and interrelate the more common operators were not given. Similarly, the notion of quantum field is, philosophically speaking, a detail belonging to the Dynamical postulate (P3). It enters the general scheme as follows. The basic types of quantum systems are classified using relativity-inspired assumptions about symmetries of the physical universe. By "basic types of quantum systems", we mean the so-called elementary particles. However, since none of them is especially particulate in the classic sense, it seems desirable to avoid the term <u>particle</u> with its classical connotations and to replace it by the appelation <u>quantum system</u>. Each type is group theoretically associated with an operator on the Hilbert space. The operator is a function of space-time, and is accordingly called a field. In a popular form of quantum field theory, these local field operators are combined to form Lagrangian operators which characterize the physical system in the

*If we formally define the Hamiltonian operator H by $T = \exp -\frac{i}{\hbar}\int H\,dt$, then $\psi(t)$ satisfies the <u>Schrödinger equation</u>, $H\psi(t) = i\hbar\frac{\partial\psi(t)}{\partial t}$.

sense that an action principle together with field commutation relations leads to dynamical equations for the field operators. Lagrangian invariance principles generate key observables as functions of the field operators; hence causal development of the observables is calculable (Heisenberg picture). In particular, this complex scheme produces an evolution operator so that P3 and related theorems come into play as stated. (In a relativistic theory, temporal development must of course always be understood relative to some given frame.)

Apparently because of its historic development as a "second quantization", modern quantum field theory is saddled with a notation which can easily mislead anyone accustomed to using only the function space representation of quantum theory, i.e., wave mechanics. In the latter, for example, an electron state vector ψ is a Dirac 4-component spinor $\psi(x,t)$ which satisfies

(1) $H_D \psi(x,t) = i\hbar \frac{\partial \psi(x,t)}{\partial t}$, where H_D is Dirac's electron Hamiltonian operator. In the field theoretic approach, as just explained, a certain spinor operator-function is associated with the electron, viz., the Dirac field, conventionally denoted by $\psi(x,t)$, a 4-component spinor operator which obeys the operator relation

(2) $H_D \psi(x,t) = i\hbar \frac{\partial \psi(x,t)}{\partial t}$.

The identical form of (1) and (2) is most unfortunate from a conceptual viewpoint; for while the $\psi(x,t)$ of (1) is a state vector characterizing measurement statistics of a certain electron ensemble, the $\psi(x,t)$ of (2) is a field operator, representing only general dynamical properties of the electron but carrying no information about the measurement statistics for any specific experiment. In short, the difference between $\psi(x,t)$ of (1) and $\psi(x,t)$ of (2) is rather like that between (the mathematical representatives of) the concepts state and system. To avoid confusion,

field theorists usually denote the state vector by a Dirac ket $|\ \rangle$; however, since the concept of quantum field is not required in general measurement theory, it suffices below to use the less cumbersome ψ to denote the vector.

3. Primitive Terms of Quantum Theory

Scanning the foregoing postulates for those physical constructs which play major roles in general quantum theory, we find seven requiring careful study: system, preparation, ensemble, observable, measurement, result (of measurement), and state. It should noticed immediately that none of these terms is intrinsically quantal; all of them have meaning, perhaps trivially in some cases, within the methodology of classical physics. However, within the quantal framework, some of them acquire extended significance and important subtleties of meaning.

Although our presentation of quantum theory mimics such rigorous mathematical systems as pure geometry by referring to primitives, postulates, and theorems, several distinctions must be recognized. When geometry is carefully axiomatized, the primitives are truly undefined; point, line, congruence, etc., are totally devoid of experiential meaning. Every relation among them is stated in the axioms, and these connections embody all properties to be associated with the terms. This information alone coupled with pure logic then leads to numerous initially hidden interrelationships among the terms, viz., the theorems. When the primitive concepts are provided with empirical counterparts via operational definitions the total scheme becomes physical geometry, the science of space. Ideally, perhaps every scientific discipline, including quantum theory, should be cast in this mathematically utopian form; but in fact even the relatively few physicists committed to logical rigor do not generally employ postulational

schemes so pure as the rather exceptional case of geometry. Unfortunately, the construction of such mathematical systems is predicated upon considerable hindsight and is therefore inapplicable in the formative stages of a theory.

Returning now to quantum theory, we do not claim that the postulates of the last section embrace every relation among the seven primitives which might be invoked while deducing their consequences, nor can we assert that the selected list of basic terms is complete. Furthermore, none of these primitive constructs will ever be regarded as absolutely undefined; and in some cases their root physical definitions to be reviewed below will later demand further qualification. In spite of these departures from mathematical propriety, it is still possible to test quantum theory for logical consistency by following a program which parallels similar considerations in more rigorous logical systems.

Because the primitive constructs mentioned above are not a priori independent and undefined, it is necessary to begin with explanations which convey their minimal physical meanings. Such accounts will suffice until tensions in the logical matrix of primitives and postulates induced by the problem of measurement create the need for further explication.

The concept of system is understood throughout physics as the actual object of study; epistemologically, it is posited as the bearer of observables and hypostatized to become said object. A mathematician might be disposed to define a system as a set of observable-symbols, but such purity misses the point. An example of a quantum system is an electron in a given environment, or equivalently, a Dirac field in a single electron state interacting with other given physical systems. Since the concept of quantum observable is philosophically more sophisticated than its classical progenitor, we postpone the details to the next section. It is sufficient

for the moment to state that whenever a system is subjected to the process called measurement of a given observable, there emerges a number, the result of the measurement. Thus observables serve to provide quantitative information about systems; every observable is endowed with measurement procedures which, if performed upon the system, yield the numerical results. Accordingly, observables are also called physical quantities. For reasons to be discussed later, we have purposely described these classically transparent concepts in what seems at first to be an overly cunbersome manner.

In accordance with the emphasis in physics on reproducibility of phenomena, a single measurement carries little significance. Systematic study of a given type of system therefore requires a well-defined, repeatable process of preparation. In general, what is of interest in physics is a set of measurement results for several observables, where the measurements are all performed upon identically prepared systems. Since acts of preparation are themselves physical processes under the governance of quantum theory, an interesting exercise related to the theory of measurement is the quantal description of a preparation. We shall return to this idea subsequently (section 19).

The collection of identically prepared systems upon which the various measurements are performed is called the ensemble; more than one philosophic stand has been taken by physicists regarding the exact status of the quantum ensemble. It turns out that the different requirements placed upon the measurement act depend strongly on different meanings attached to the ensemble concept and the related construct, physical state. A later section will contrast the various kinds of ensembles used in physics in order to identify the rather unique character of the quantum ensemble.

In all physical theories, the state of a system refers to its momentary physical condition; it is the seat of causality in physics in the

sense that some law of motion controls its temporal evolution. By "physical condition" is meant that states are related somehow to observables, and to measurement results. In the form given above the quantal postulates seem to correlate the state concept to a system only through an intervening ensemble of such systems identically prepared. Only the statistics of measurement results obey a causal law. Thus, in effect quantum theory seems to shift the reference of the state concept from the single system to the ensemble.

It might be objected that this modification is illusory, that the postulates were stated with distorted emphasis on ensembles which hides the true meaning of state. Thus, classical statistical mechanics might be axiomatized in a similar format; but the classical individual state would be lurking in the shadows and could be exposed with the proper logical illumination. Elsewhere[4] we have carefully examined this question and demonstrated that, while such is indeed the case for classical statistics, no such analogous reduction to individual states is possible within the quantum framework.

A quantum state refers to an ensemble; an ensemble is defined by its mode of preparation and characterized by the statistics of measurements performed upon its member systems, and these statistics determine the state. Thus it is often convenient to speak of a preparation of state, a concept emphasized by Margenau[8] to delineate a class of physical processes often erroneously called measurements. This completes our preliminary survey of the key terms of general quantum theory; but before attempting to describe quantally the process of measurement, we shall undertake deeper analysis of the constructs observable and ensemble.

4. The Nature of Quantum Observables

To those physicists who take mathematics to be in the same category as metal-working lathes and vacuum pumps, the subtleties taught by modern mathematicians often seem inane. Among these is the difference between a function f and its range value at domain point x, f(x). Even in classical mechanics, however, there are two instances in which this distinction is physically meaningful, for it represents a philosophically important dichotomy among physical constructs. Consider first the case of the Hamiltonian H, a function of phase (q,p). Here the value of logically distinguishing H and H(q,p) is eventually recognized by anyone thoughtfully studying analytical mechanics. In fact, the term functional form is often used to stress that the function itself, not its value, is under consideration. H itself is the mathematical representative of the system of interest; i.e., "the functional form of H(q,p)" contains the dynamical characteristics of the system and represents it in the law of motion. The numerical value of H(q,p), on the other hand, is usually just the result which would be obtained if an energy measurement were performed on the system. (For the sake of familiarization, we continue to use this tedious phraseology introduced earlier for minimal descriptions of measurement.) Failure to take note of the difference between H and H(q,p) can actually lead to faulty reasoning of physical significance. For example, consider a mechanical problem with initial conditions $H(q,p) = E$ and $q = q_0$. (E and q_0 are constants.) Hamilton's equations, used properly, will determine the motion. But consider the following reasoning: since H(q,p) is not explicitly time dependent, $H(q,p) = E$ not just initially but throughout the motion, according to a basic theorem. Hamilton's equations therefore become especially simple if the constant E is substituted for H. The immediate results are $\dot{p} = -\frac{\partial H}{\partial q} = -\frac{\partial E}{\partial q} = 0$, $\dot{q} = \frac{\partial H}{\partial p} = \frac{\partial E}{\partial p} = 0$;

hence $q(t)=q_0, p(t)=0$. The particle just sits still! That this solution is wrong is easily seen by considering $H=\frac{p^2}{2m}+\frac{R}{2}q^2$, the simple harmonic oscillator, which under initial conditions of the type given is not in general immobile. The error lay of course in neglecting the distinction between a function and its value.

There is a second instance in classical mechanics where this methematical point could be stressed; it was not, however, until the advent of quantum mechanics that its message became apparent. Because of its dynamical significance, the function H is rather special; but in classical mechanics, every function of phase has physical meaning. A function corresponds to an observable; and the value of a function for a state (q,p) is, again in our "minimal" phraseology, the result which would be obtained if a measurement of the observable were performed upon a system in said state. Since the state uniquely determines the measurement result for every observable through the corresponding function, the natural classical manner of describing the situation was not a minimal account but rather the simpler assertion that in a state (q,p) the system had an observable A of value A(q,p). For example, "the oscillator has an energy of 30 ergs"; and, of course, if an energy measurement is performed, the result would be 30 ergs --but to state this explicitly seems pointless. Thus with the notable exception of H, classical mechanics did not require rending the function from its value, nor the observable from its measurement result; and the concept of measurement entered only implicitly into physical discourse.

A glance at the postulates and theorems in section 2 shows that no such departure from the minimal terminology is admissible in quantum theory. There the constructs observable and measurement result are related only via probabilistic connections, and measurement thereby emerges as a construct which must appear explicitly in quantal propositions.

This separation of the concepts observable and measurement result is of considerable importance to the philosophic understanding of quantum physics. The peculiar nature of quantal observables has been depicted in several ways, three of which we shall briefly review: Bohr's complementarity principle,[9] Margenau's latency theory,[10] and Heisenberg's "potentia" doctrine.[11]*

Complementarity is accorded at least token recognition in virtually every introductory quantum text. Its basic premise is apparently that the nature and results of microphysical research demonstrate that what we called a minimal account is also a maximal account. The utter impossibility of direct perception of atomic objects suggests the separation of observable and measurement result; the failure of all attempts to preassign unique measurement results to all observables by careful preparation of state renders the separation final. Thus, given an atomic object (including a mode of preparation), there is the choice of measuring any observable and a theory which provides probabilities for the possible results; but to say that the system has position q_o, momentum p_o, energy E_o, etc. is physically meaningless. Accordingly, Bohr coined the term complementary to describe this characteristically quantum relationship among the observables.

Margenau's latency theory classifies observables by the terms possessed and latent. In the classical proposition that a particle has a certain energy, the energy is clearly understood as a property possessed by the system. Similarly, any classical function of state--mechanical, electrodynamic, or thermodynamic--denotes an observable attached possessively to a system. Nevertheless, in classical physics there are also

---

*Outside the present context, the divergence of these three viewpoints far exceeds their similarity, as later sections will illustrate.

observables associated with systems not possessively but responsively; i.e., if subjected to a certain environment, a system displays a property not constantly exhibited. For example, the acoustic "observable" pitch is inapplicable to a vibrating reed in an evacuated box; but if the enclosure is opened to the atmosphere, a "value" of the pitch emerges. Such is the nature of a latent observable. Because it is impossible to assign values to all the observables of a quantum system in a possessive way and because quantum theory unavoidably deals only with statistics of measurement results, most of its observables are latent. It is strictly improper to speak of a quantum system's having energy E_o; the strongest admissible statement is the conditional one that, if an energy measurement were performed, the result would, with some calculable probability, be E_o. Quantum observables are thus latent in the sense that their values appear only in response to measurement. Quantal latency for a given observable is represented mathematically by the existence of physically realizable state vectors which are not eigenvectors of the corresponding operator, i.e., states for which measurement results for the observable in question are irreducibly unpredictable. Hence, almost all quantal operators correspond to latent observables.* Currently one exception is the mass, a defining parameter for a given type of quantum system, but conceivably an ultimate quantum theory might deal with a single kind of system for which all the "elementary particles" are but states belonging to the eigenvalues in the mass spectrum. In such a theory mass, too, would be a latent observable (provided superpositions of different mass eigenvectors were physically realizable).

---

*Perhaps those which generate superselection rules are an exception: in one form, superselection rules exclude pure states which are not eigenstates of certain observables (cf. sec. 19); therefore it would always be possible to regard such observables as possessed.

Another description of the nature of quantum observables appears in Heisenberg's discussions of the Copenhagen interpretation. The state of a quantum system before measurement is envisaged as a set of tendencies likened to Aristotelian potentia. Upon measurement one possibility is fulfilled, as an actual, perceptible event occurs, culminating in the extraction of a number. Measurement of an observable is thus depicted as a "transition from the possible to the actual", which is but another way to state the latent character of the construct observable in quantum theory.

5. The Nature and Purpose of Ensembles

In any physical theory which assigns probabilities to possible measurement results, use of the construct ensemble is unavoidable, simply because probability in physics means relative frequency. This is not to say, in sterile operationist fashion, that through this empirical definition probability acquires its total significance. The situation in probability theory is no different in this respect from the rest of physics; i.e., constructs are endowed not only with empirical definitions but also with theoretical ones. In the case of probability, the theoretical side has long been in controversy; rival mathematical schools lay the foundations in different ways. As with other mathematical choices, physicists adopt the most naively intuitable version capable of meeting their needs. However, in any case probability, as drawn from mathematics, is logically a primitive term defined implicitly by the axioms in which it is embedded. It obtains physical meaning only when the rule of correspondence[12] is invoked which correlates it to the relative frequency of measurement results.

Although the foregoing remarks correctly portray the epistemological status of physical probability, the historical development of course did not proceed so logically. As Carnap[13] has noted, the search for a good

theoretical definition of probability is a problem of explication, i.e., replacement of an old, vague concept by a new, exact one. Thus the logically prior mathematical theory was itself inspired by common-sense notions of probability as a measure of tendencies or propensities for events to occur. Undoubtedly such ideas likewise underlie--and perhaps undermine--the physicist's conception of probability.

Consider, for example, the following proposition: if a measurement of observable A is performed upon a system (prepared in a specified manner), the probability of obtaining the result a_i is W_i. Superficially, layman and physicist alike construe W_i as somehow reflecting a tendency for the emergence of a_i from that system at the instant of measurement; but careful consideration reveals the rather mystical tenor of that view. Actually W_i is a clearly defined quantity, viz., the relative frequency of the result a_i arising from A-measurements upon an ensemble of identical systems S all prepared in the manner Π. Whatever physical information about the pair (S, Π) probabilities like W_i may carry, the connection between system and probability is always through the intermediary construct, ensemble; otherwise, probability is a concept too hazy to qualify for a place in physics.

So far we have discussed only the precise meaning of physical probability; that analysis now justifies a shift in emphasis from probability itself to the intimately related notion of ensemble, since a probability without an ensemble is unphysical. Of special importance, the various ways in which probabilities enter theoretical physics are mirrored in the nature and purpose of the associated ensembles. The relevance of a study of classical and quantal ensembles to the quantum theory of measurement will become clear in later sections.

A physical ensemble is basically a set of identically prepared

independent systems. In principle, measurements can be performed on the constituents at any instant after preparation; and the set is sufficiently large to warrant statistical analysis of the measurement results, including meaningful identification of probability as relative frequency. However, the phrase, "a set of identically prepared systems", represents an abstraction physically realizable in several ways. The term "set" denotes a mental collection of objects which need not even coexist; a set may be an aggregation of elements all present at once, a temporal sequence of single elements, or any admixture of these two extremes. Similarly, "identically prepared systems" might refer to just one system prepared and sequentially reprepared. Whichever combination is selected, the member systems are strictly independent; for example, the assemblage of molecules constituting a real gas is not an ensemble of molecules.

The physical significance of an ensemble depends not just on its structure but also on its purpose, i.e., on the connection between the ensemble and the actual physical situation to which it refers. In the classical realm, perhaps the simplest ensemble imaginable is a collection of coexisting, noninteracting mechanical systems. If their common preparation process consists of placing a system in a given dynamical state, the resultant ensemble will then be homogeneous, or pure, for obviously every subensemble is identical to the whole ensemble so far as measurement statistics are concerned. However, if the mode of preparation is less discriminating, there will be a distribution of states over the ensemble, sets of measurement statistics will vary among subensembles, and hence the whole ensemble will be mixed. Although this simple ensemble thus illuminates the basic physical meaning of ensemble homogeneity, defined mathematically in section 2, it leaves the impression that the homogeneity concept is all too trivial to be of any value. However, this seeming triviality is but a

manifestation of the classical context in which the example was given; in particular, it was implicit in the language used that the observables, hence the classical states, were possessed, a property which enables a convenient pictorial conception of the systems.

In Gibbsian statistical mechanics, an ensemble of the type just described is employed; but it is not used directly, i.e., the physical system of interest is not itself a collection of coexisting, noninteracting, identically prepared systems. In fact, it is just one such system, related to the imaginary ensemble of replicas by a postulated correspondence between observed values and ensemble averages. Why, then, is an ensemble used at all? From a strict mechanistic viewpoint, the reason might be simply that thermodynamic systems, whose behavior Gibbs sought to comprehend mechanically, are incredibly complex. Actual knowledge of a precise mechanical state for the septillion molecules in a mole of gas is a practical chimera. Gibbs' virtual ensemble could be regarded, therefore, as a mathematical representation of such ignorance. In fact, the scheme was later generalized to become modern information theory. It should be stressed, however, that we have asserted only that the Gibbsian ensemble permits consistent interpretation in terms of ignorance, not that it must be so understood. Indeed, so long as the physical significance of ergodic theory remains in dispute, there is a possibility that even a complete mechanical state specification of a complex system would not account for its thermodynamic behavior, in which case the Gibbsian ensemble would be a physical construct far more abstract and fundamental than its "ignorance interpretation" suggests. Gibbs' ensemble allows the ignorance interpretation chiefly because it is framed within a classical metaphysic which provides something to be ignorant of, viz., the values of possessed observables.

Consider now the transition to the latent observables of quantum

theory and its impact upon everything said about classical ensembles. That simple ensemble consisting of a simultaneous assembly of noninteracting systems identically prepared now requires a "minimal" description. Strict emphasis on measurement results and the correlation of their statistics to modes of preparation replaces the graphic account in terms of individual classical states. The first preparation instruction given above, "place each system in a given dynamical state" is now quite meaningless. In view of the latency of quantal observables, the most that can be said is to "use a method of preparation such that the associated statistics of measurement results indicate a homogeneous ensemble". Similarly, some preparation schemes will produce ensembles whose measurement statistics are summarized by an inhomogeneous mean value functional. The essential point is that this latency-enforced revision of ideas destroys the basis for interpreting ensembles as expressive of ignorance in the Gibbsian sense, for in quantum physics there are no longer even in principle any innate quantities of which to be "ignorant". In quantum theory the actual object of study is effectively the ensemble itself; however, that ensemble may be any of the types described earlier in this section. In particular, it might even be a single quantum system in a temporal alternating sequence of identical preparations and diverse measurement operations.[8]

As mentioned earlier, we have demonstrated elsewhere that, by contrast to the classical case, in quantum theory the concept of homogeneity cannot be used in a consistent way to assign physical states to single elements of an ensemble. Although ordinary physical jargon speaks of "a system in the state ψ", that phrase can only mean either (1) an element of a pure ensemble with density operator $\rho = P_\psi$, or (2) an element of a pure subensemble $(\rho^{(i)})$ of a general mixed ensemble whose density operator may be expanded as $\rho = w_1\rho^{(1)} + w_2\rho^{(2)}, w_1 + w_2 = 1, w_1 > 0, w_2 > 0, \rho^{(1)} = P_\psi$.

The fact that in case (2) the same element may equally well be called "a system in (another) state ϕ " (since the expansion of ρ into pure subensembles is not unique) proves the absurdity of literally associating a state vector with a single member of an ensemble. (For elaboration, see reference 4.)

Nevertheless, because any mixed ensemble can in principle be subdivided into sets of pure subensembles, there is a logically weak sense in which the quantum mixture is often linked to ignorance: by analogy to the Gibbsian case, the mixed ensemble is sometimes interpreted to mean that there is ignorance as to which pure state the system is "really in". Indeed in the discipline called quantum statistical mechanics, this fiction is artificially upheld by conjuring up "two averages" from the quantal mean value expression $m(A) = Tr(\rho A)$. Suppose $\rho = \sum_k w_k P_{\psi_k}$ is one among the many ways the ensemble at hand can be grouped into pure subensembles, w_k being the fraction of the original ensemble in the P_{ψ_k}-subensemble, if this particular selection is made. Then, $m(A) = Tr(\rho A) = \sum_k w_k \langle \psi_k, A\psi_k \rangle$ is the average result of a-measurements on the ensemble with density operator ρ . Now, despite the fact that this expansion is not unique, it is standard practice in statistical mechanics to declare, as for example ter Haar[14] does, that m(A) "is twice an average. First we take the quantum mechanical average...in a system described by the wave function ψ_k , and, secondly, we take the average over the ensemble." The introduction to the same chapter typically cautions that in quantum ensemble theory, "one must be extremely careful to make a clear distinction between the statistical aspects inherent in quantum mechanics and the statistical aspects introduced by the ensembles". Such statements sound as though (1) ψ_k -statistics are not related to ensembles, and (2) that mixed density operators always refer to ensembles made up of systems

"really in" pure states ψ_R .

Actually, (1) and (2) are both false; however, if we make the spurious identification of ψ_R as the quantal analogue to a classical state (as noted above, there is none) and interpret "ensemble" in (2) as meaning "virtual assembly of coexisting, identically prepared systems", there results a pseudo-analogue to Gibbs' method which is of motivational value in quantum statistical mechanics. To make the analogy complete, there must be a postulated connection between $Tr(\rho A)$ and observed values of thermodynamic quantities. Such a postulate together with knowledge of specific ρ's is essentially the logical core of quantum statistical mechanics; meaningless "classical" analysis of quantum ensembles is not necessary, although it may serve to <u>suggest</u> the formulation of useful ρ's. However, its intuitive value in this context should not be mistaken for rational physics.

This digression on quantum statistical mechanics was not made to condemn its heuristic methods but to repudiate the erroneous idea that the general density operator represents ignorance in perfect analogy to the Gibbsian model; the density operator is not at all the sole property of quantum statistical mechanics but is actually a basic quantal construct. In fact, a mixed ρ cannot refer to an ensemble of systems each "really in" a pure state since, as we have repeatedly emphasized, that phraseology is logically ambiguous. A mathematically parallel situation in classical optics arises for polarization of light. If a light beam is, for example, unpolarized, we cannot meaningfully conclude that there are "really" two incoherent "sub-beams" of equal weight each linearly polarized but along perpendicular directions, for the analysis is not unique. With equal justification, many other such dissections of the unpolarized beam may be performed, among these the assertion that the "sub-beams" are "actually"

circularly polarized in opposite senses. Empirically, every unpolarized beam can be split either way with the proper equipment; thus propositions about the "hidden structure" of the unpolarized beam are physically meaningless. Just as there are light beams which are intrinsically unpolarized or partially polarized, there are quantum ensembles which are intrinsically mixed; neither has anything to do with ignorance.

To complete this discussion of ensembles, we draw attention to a striking difference between the classical and quantal cases. A classical ensemble is described by the set of probabilities that member systems are in the various pure subensembles. Because the latter correspond to classical states, they are not only statistically homogeneous (as defined in section 2) but also dispersionless, which means that for any observable, measurement results from a pure subensemble are all identical. The collection of pure subensembles into which any given ensemble may be resolved is unique.[4] Because it is physically possible to select the unique, pure, dispersionless subensembles from the original ensemble, the above mentioned probabilities may be called reducible, a term used in this connection by Margenau.[15]

For a quantum ensemble, the reduction to pure subensembles is no longer unique;[16] nevertheless, similar selection processes are still possible. Once a resolution to homogeneous subensembles has been specified, the total ensemble is then characterized in part by reducible probabilities just as in the classical case. However, by a theorem of von Neumann,[17] no quantum ensemble, not even a homogeneous one, is dispersionless. Therefore, reduction to pure subensembles does not explain away all probabilities; there always remain probabilities, called irreducible by Margenau,[15] which reflect the intrinsic dispersion of homogeneous quantum ensembles. Incidentally, this property is the backbone of Heisenberg's Principle of

Indeterminacy. To summarize: a classical ensemble may be reduced to a unique set of homogeneous, dispersionless subensembles; a quantum ensemble may be reduced to any one of numerous sets of homogeneous subensembles each of which invariably exhibits dispersion in the statistics of measurement results for most observables.

6. Outline of Standard Measurement Theory

To provide a skeletal basis for subsequent discussions of its many ramifications, the "standard" quantum theory of measurement[18] will now be outlined in an abstract mathematical fashion temporarily avoiding all philosophical problems of interpretation. Any actual measurement of an observable a on a quantum system S assumed to be an element of a known ensemble, is performed with an auxiliary system $M(a)$ called a measuring apparatus for observable a, or a-meter. This means that S and M physically interact so that known correlations arise between the possible measurement results of observable a and some observable c belonging to M. Since M is an a-meter, these correlations are sufficient to render a "direct" a-measurement superfluous. Thus a is measured by "reading the a-meter", i.e., by measuring c on M.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the Hilbert spaces associated with S and M, respectively. The tensor product space $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$ is then appropriate for the study of the $S-M$ interaction. As usual, the operators corresponding to a and c will be denoted by $A\otimes 1$ and $1\otimes C$. To avoid burdensome notation, we assume for the present that A and C have discrete, nondegenerate spectra. A, C satisfy $A\alpha_R=a_R\alpha_R$, $C\gamma_c=c_c\gamma_c$; $\{\alpha_R\}$ spans $\mathcal{H}_1$, $\{\gamma_c\}$ spans $\mathcal{H}_2$, and $\{\alpha_R\otimes\gamma_c\}$ spans $\mathcal{H}$.

Nothing significant comes from considering mixtures as opposed to

pure states; we assume therefore that initially S and M are "in" pure states ψ and χ. (Having made the point that the ensemble must not be forgotten, we shall henceforth often use this common expression.) It can be shown that the composite system $S+M$ will then be in state $\psi\otimes\chi$

By Th9 the temporal evolution of the state vector is always expressible by a linear evolution operator T: $\psi(t_2)=T(t_2,t_1)\psi(t_1)$. In the product space $\mathcal{H}$, if S and M do not interact, T is decomposable to $T_1\otimes T_2$; conversely, an indecomposable T expresses interaction.

Now, according to the general principles stated above, the measurement process entails $S-M$ interaction leading to the establishment of correlations. Mathematically, this will be expressed as a condition on T_A, the indecomposable evolution operator for a-measurement. That condition, which we call the correlation assumption, is almost always given as

$$T_A(\alpha_R\otimes\chi)=\alpha_R\otimes\gamma_R$$

from which it follows that

$$T_A(\psi\otimes\chi)=T_A\left(\sum_R\langle\alpha_R,\psi\rangle\alpha_R\otimes\chi\right)=\sum_R\langle\alpha_R,\psi\rangle\alpha_R\otimes\gamma_R.$$

The desired correlation arises as follows: it can be shown from the axioms that the final composite state vector $\sum_R\langle\alpha_R,\psi\rangle\alpha_R\otimes\gamma_R$ means that, if an a-measurement is performed on S and a c-measurement on M, the probability that the pair (a_R,c_ℓ) will result is just $|\langle\alpha_R,\psi\rangle|^2\delta_{R\ell}$. Hence, the c-measurement alone suffices.

Finally, it is customary to consider the post-measurement S-ensemble independently; this is done by focusing on the measurement statistics for S-observables only,[19] i.e., those corresponding to operators of the form $B\otimes 1$. A simple calculation shows that the density operator $\hat{\rho}_1$ of that ensemble is given by* $\hat{\rho}_1=Tr_2\,P_{\sum_R\langle\alpha_R,\psi\rangle\alpha_R\otimes\gamma_R}=\sum_R|\langle\alpha_R,\psi\rangle|^2P_{\alpha_R}$.

* Tr_2 signifies a trace operation involving only matrix indices representing $\mathcal{H}_2$; for elaboration on this procedure for finding the density operator for a subsystem, see Ref. 4, sec. 8.

The structure of $\hat{\rho}$ immediately indicates one of the ways in which the S-ensemble after a-measurement is reducible to pure subensembles. Each subensemble in this method of selection is dispersionless relative to a-measurements since its state vector is an eigenvector of A. Moreover, the subensemble for which an a-measurement certainly yields a_k is just the fraction $w_k = |\langle \alpha_k, \psi \rangle|^2$ of the original ensemble, in perfect harmony with the basic quantum theorem (Th3, section 2) that a-measurements on systems in state $\psi = \sum_k \langle \alpha_k, \psi \rangle \alpha_k$ yield a_k with probability

$$W_a(a_k; \rho = P_\psi) = Tr(P_\psi P_{\alpha_k}) = |\langle \alpha_k, \psi \rangle|^2.$$

This completes our review of the mathematical skeleton of quantum measurement theory. Ensuing sections will explore many controversial facets of the theory which were purposely glossed over in this preliminary outline. Such a critical survey will lead eventually to a clarification of the meaning and epistemological status of the two primitive constructs not yet fully analyzed--measurement and preparation.

7. The "Optics" of Measurement

The preceding mathematical orientation to measurement theory cannot be taken as the unique core of all philosophical and theoretical discussions on the subject. As will become increasingly evident, there are many variations on that principal mathematical theme. However, one key point is accorded almost universal acceptance as the fundamental desideratum of a quantum theory of measurement, viz., the proposition that a-measurements upon an ensemble whose initial density operator is $\rho^{(n)} = P_{\psi_n}$, $\psi_n = \sum_k \langle \alpha_k, \psi_n \rangle \alpha_k$, where $\{\alpha_k\}$ are the eigenvectors of a's operator A, will produce a post-measurement ensemble whose density operator is

$$\hat{\rho}^{(n)} = \sum_k |\langle \alpha_k, \psi_n \rangle|^2 P_{\alpha_k}.$$

From this it follows that a-measurements upon an initially mixed ensemble

characterized by $\rho = \sum_n w_n \rho^{(n)}$ induce this transformation in the density operator: $\rho = \sum_n w_n \rho^{(n)} \longrightarrow \hat{\rho} = \sum_n w_n \hat{\rho}^{(n)}$ $= \sum_n w_n \sum_{\alpha_R} \langle \alpha_R, P_{\psi_n} \alpha_R \rangle P_{\alpha_R}$. This may be expressed as $\rho \longrightarrow \hat{\rho} = \sum_{\alpha_R} \langle \alpha_R, \rho \alpha_R \rangle P_{\alpha_R}$, the form used by von Neumann;[20] or if P_{α_R} is written $|\alpha_R\rangle\langle\alpha_R|$ (Dirac notation), this "a-measurement transformation" assumes another simple form, $\rho \to \hat{\rho} = \sum_{\alpha_R} P_{\alpha_R} \rho P_{\alpha_R}$. Henceforth we shall frequently refer to $\rho \to \hat{\rho}$ simply as <u>the</u> (von Neumann) <u>measurement transformation</u>.

Thus the philosophical challenge of the measurement concept in quantum theory is generally translated into mathematical physics as follows: prove that the measurement interaction of a system with an a-meter transforms the system density operator in the manner just defined. Sometimes this problem is expressed in the colorful language of waves:[21] prove that measurement "destroys coherence" or "introduces random phase relations". The origin of these phrases is to be found in the historic analogy between quantum mechanics and classical optics, which is especially clear in the familiar Schrödinger wave mechanics. In that analogy the pure state $\psi = \sum_{\alpha_R} \langle \alpha_R, \psi \rangle \alpha_R$ corresponds to "white light", a coherent superposition of the various "colors" $\{\alpha_R\}$ which "interfere" in the sense that mean values generally involve cross terms. For example,

$$m(B) = Tr(P_\psi B) = \sum_{\alpha_R} |\langle \alpha_R, \psi \rangle|^2 \langle \alpha_R, B\alpha_R \rangle + \sum_{\alpha_R \neq \alpha_\ell} \langle \psi, \alpha_\ell \rangle \langle \alpha_R, \psi \rangle \langle \alpha_\ell, B\alpha_R \rangle$$

A measurement interaction is then depicted somewhat like passage of light through a prism which separates the various "colors" so that there is no longer any "interference" among them, hence no cross terms:

$$\hat{m}(B) = \sum_{\alpha_R} |\langle \alpha_R, \psi \rangle|^2 \langle \alpha_R, B\alpha_R \rangle.$$

The last equation may be written as $\hat{m}(B) = Tr(\hat{\rho} B)$, where $\hat{\rho} = \sum_{\alpha_R} |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}$; therefore, an a-measurement is said to remove the "interference of probabilities" in the "coherent superposition"

$\psi = \sum_R \langle \alpha_R, \psi \rangle \alpha_R$ by a transformation to the "incoherent superposition" $\hat{\rho} = \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}$.

Another imaginative portrayal of the effects of measurement rests on the contention that measurement introduces "random, uncontrollable phase relations". Thus if $\psi = \sum_R \langle \alpha_R, \psi \rangle \alpha_R$ becomes $\hat{\psi} = \sum_R \langle \alpha_R, \psi \rangle e^{i f_R} \alpha_R$ where $\{f_R\}$ is a set of "random" phases to be averaged out, then effectively the same transformation is obtained, since the average is calculated as follows:

$$\hat{m}(B) = \prod_n \int_0^{2\pi} \frac{df_n}{2\pi} \langle \hat{\psi}, B\hat{\psi} \rangle = \sum_R |\langle \alpha_R, \psi \rangle|^2 \langle \alpha_R, B\alpha_R \rangle + \sum_{R \neq \ell} \langle \psi, \alpha_\ell \rangle \langle \alpha_R, \psi \rangle \langle \alpha_\ell, B\alpha_R \rangle \frac{1}{(2\pi)^2} \int_0^{2\pi} df_\ell \int_0^{2\pi} df_R \, e^{i(f_\ell - f_R)}$$

$$= \sum_R |\langle \alpha_R, \psi \rangle|^2 \langle \alpha_R, B\alpha_R \rangle$$

$$= \mathrm{Tr}(\hat{\rho} B), \quad \hat{\rho} = \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}.$$

This phase idea is reminiscent of the old description of an unpolarized light beam as two beams linearly polarized in perpendicular directions but having "random phase relations" between them; but it is noteworthy that classically this described only an algebraic trick, not the actual nature of unpolarized light.

Unfortunately, such picturesque analogies to classical optics can never deepen our understanding of quantum measurement; indeed, they may even becloud the real issues. Although it is undeniable that quantum theoretical <u>calculations</u> often bear striking resemblance to those of classical optics and acoustics (for obvious historical reasons) and that working physicists therefore draw heavily on such <u>mathematical</u> parallels in the course of everyday problem solving, nevertheless, as <u>physical</u> theories, wave optics and "wave" mechanics are strikingly distinct. As a consequence of this physical gulf, these mathematical analogies may offer more confusion than illumination when applied to a problem so fundamental as the quantum theory of measurement.

Regardless of the description given the measurement transformation $\rho \rightarrow \hat{\rho}$, a basic question remains: why is this transformation so often assumed as the goal of measurement theory? What properties does it have so essential to the mathematization of the measurement process?

Those who have developed an intuition for the optical analogy are supposedly able to induce the principle of coherence destruction from idealized "electron interference" experiments, for which the empirical warrant is the related work of Davisson[22] and Germer. An electron is prepared for study by having an electron gun fire it at a barrier impenetrable but for two slits; various measurements made on the side of the slits opposite the gun are then considered. A simple position measurement can be made by placing a fluorescent screen parallel to the two-slit wall; the co-ordinates of the glowing dot where the election is absorbed are the results of the position measurement. It is well known that the arrangement of dots associated with an ensemble of electrons identically prepared and measured simulates the optical interference pattern of Young's two-slit experiment. However, if two devices capable of detecting the passage of electrons through them are interposed between slits and screen adjacent to the two slits, it becomes possible to measure the position not only at the fluorescent screen but also at the slits. Now, the array of dots on the screen caused by electrons after passage through these new detectors is not at all like a Young interference pattern but resembles instead two overlapping one slit diffraction patterns. It is therefore tempting to conclude that "destruction of interference" is an essential characteristic of the measuring process, since the measurements at the slits have just that effect.

However, two cautionary remarks are in order concerning this argument: (1) Analysis of a single gedankenexperiment (or a single class of actual

experiments) is obviously insufficient to establish any universal property of measurements; the quantum theory of measurement should not be founded on specialized empirical knowledge. (2) We have here the first indication of conceptual interlocking among the notions preparation, interaction, and measurement--a theme to be developed further in later sections. There is good reason to question the logical appropriateness of saying in the above gedankenexperiment that the position measurement at the slits affected the coherence; indeed, even if no number emerged (hence no measurement), the interaction with the detectors would have destroyed the interference anyhow. In view of this physically obvious fact, is it not perhaps more reasonable to interpret the foregoing discussion of electron interference primarily as illustrative of the concept preparation rather than measurement? Let ψ_1, ψ_2 be the state vectors associated with preparations involving just one slit with no adjacent detector. Roughly speaking, the gedankenexperiment actually proved only this: if there are two slits without adjacent detectors, the apparatus prepares a pure ensemble with $\rho = P_{c_1\psi_1 + c_2\psi_2}$; if the two slits have adjacent detectors, the total apparatus prepares a mixed ensemble with $\rho = w_1 P_{\psi_1} + w_2 P_{\psi_2}$. The conversion of the pure ensemble to the mixture is fully explicable in terms of its interaction with the detectors, whether the latter are used to perform measurements or not. Thus such a demonstration does not necessarily reveal any significant feature of measurement in general.

The literature of quantum theory abounds with mysterious interpretations of this electron interference gedankenexperiment; for example, the presence and absence of the detectors at the slits has been said to reveal the "complementary", "dual" natures of the electron as particle and wave, respectively. We prefer to circumvent such terminology by regarding the electron always as simply a quantum system and never as either particle or

wave. (The axioms of section 2 referred to neither classical construct.)

In the context of measurement theory, preoccupation with the "undulatory aspect of the electron" leads to an especially mystical derivation of the measurement transformation. The common jargon of quantum physics, if taken literally, welds the electron too tightly to the wave function (state vector); thus the phrase "an electron with wave function ψ" suggests more the false picture of some amorphous undulation accompanying the electron than the correct meaning of ψ in terms of ensemble measurement statistics relative to a given preparation. In terms of these strange individual electron waves, the interference gedankenexperiment may be explained as follows: when no detectors are adjacent to the slits, every electron wave has the form $\psi = c_1\psi_1 + c_2\psi_2$; but when detectors are included so that measurements are performed, the measurement act introduces random, unpredictable, hence uncontrollable, phase factors into the individual electron waves. These random phases are averaged out in computing mean values, a prodecure equivalent to describing the post-measurement ensemble by the mixture $\hat{\rho} = |c_1|^2 P_{\psi_1} + |c_2|^2 P_{\psi_2}$, as we showed earlier in this section. Thus the introduction of random phases is sometimes taken as an essential property of measurement. Needless to say, within the quantal framework developed in previous sections, the foregoing derivation of this random phase principle is logically absurd; as a matter of fact, because of its untenable association of waves (state vectors) with single elements of the ensemble, the proposition cannot even be rigorously stated.

This survey of the "optics" of measurement has not produced any convincing answer to the question posed earlier: why is proof of the transformation $\rho = P_\psi \rightarrow \hat{\rho} = \sum_R |\langle q_R, \psi\rangle|^2 P_{q_R}$ so commonly accepted as the goal of the quantum theory of measurement? Since the popularity of this view is undoubtedly due to its compatibility with the Copenhagen

interpretation of quantum theory, we turn now to that philosophy for a deeper explanation.

8. Measurement in the Copenhagen Interpretation

Because of the explicitness of the proposition which we are asking the Copenhagen school to justify, general epistemological considerations like those of Bohr and von Weizsäcker are not too helpful. On the other hand, Heisenberg tends to be more specific in his philosophic discussions and has in fact given detailed expositions of the nature of measurement; accordingly, we shall take him as spokesman for the so-called* Copenhagen version of quantum measurement theory.

Consider, as Heisenberg did in his early book[23] on quantum theory, a beam of atoms, "all of which are initially in the state n" (i.e., a pure ensemble with state vector ψ_n, an energy eigenvector) sent through a force field inhomogeneous in the direction of the beam. If energy measurements are performed on atoms of the beam emerging from this field, the measurement results will disperse, the relative frequency of the result E_n being given by $|S'_{nm}|^2$, defined by $\phi_1 = \sum_k S'_{nk} \psi_k$, where ϕ_1 is the state vector into which the initial ψ_n causally developed due to interaction with the field. Heisenberg uses slightly different terminology: "...$|S'_{nm}|^2$ is the probability of finding an atom in the state m after it has emerged from the field..."[23]; and the latter is said "to cause transitions to other states".[23]** If such measurements are not made and the

---

*It might well be argued that there are many "Copenhagen interpretations" and that the present section deals with the Heisenberg "Copenhagen interpretation" as opposed, for example, to the Bohr "Copenhagen interpretation"; however, we shall not enter into that debate. In the present context, the term "Copenhagen interpretation" will be used in the same way Heisenberg uses it.

**None of the italics in the quotations of this section are in Heisenberg's original papers.

beam is once more sent through a similar field, the state vector develops from ϕ_1 to $\phi_2 = \sum_m S'_{nm} \sum_\ell S''_{m\ell} \psi_\ell = \sum_\ell (\sum_m S'_{nm} S''_{m\ell}) \psi_\ell$, so that energy measurements now yield E_ℓ with relative frequency $|\sum_m S'_{nm} S''_{m\ell}|^2$, a situation again expressible in the language of "transitions".

So far, the differences between Heisenberg's description and our minimal account seem entirely semantic; but this is not the case, for Heisenberg goes on to make further generalizations which are deducible only from his version. Suppose that between the fields, "the atoms...are disturbed by the performance of an experiment which would have made possible the determination of the stationary state. The result of the experiment is not observed, however. The probability of the state ℓ is then $\sum_m |S'_{nm}|^2 |S''_{m\ell}|^2$ ".[24] Mathematically, this is equivalent to the statement that the between-the-fields energy measuring device brings about the measurement transformation introduced in the preceding section. Thus Heisenberg effectively endorsed the claim that this transformation represents a universal trait of measurement processes. Unfortunately, the only justification he offered at the time was a semiclassical (hence unconvincing) analysis of measurement-induced phase uncertainties in deBroglie waves.[25] However, better arguments do appear in his later writings and will be presented below, but first a final general conclusion supposedly illustrated by the atom beam must be noted.

The expression $\sum_m |S'_{nm}|^2 |S''_{m\ell}|^2$ was said to be the "probability of the state ℓ " provided an energy measurement operation was performed between the fields without observing the result. From the general perspective on quantum theory advocated in our opening sections, such a proviso seems utterly irrelevant; the relative frequency of a given measurement result obtained from atoms prepared by passage through two fields

with some apparatus between them can depend only on the physical nature of that apparatus--not on its purpose. Yet Heisenberg says that if the result of the between-the-fields energy measurement is observed, then "the atom is known to have been in state m [between the fields]. The probability of the state ℓ is then given by $|S''_{m\ell}|^2$."[24] This manner of speaking is unfortunate, for it suggests that physical probability is contingent upon an observer's knowledge. To avoid that conclusion, we are compelled to interpret Heisenberg's last proposition as follows: if the original ensemble is truncated by measuring after passage through the second field only those atoms which yielded E_m at the energy measurement between the fields, then the new probability for the final result E_ℓ is $|S''_{m\ell}|^2$. This is equivalent to attributing to the measurement process the property that the subensemble of atoms which yielded E_m has as its post-measurement density operator P_{ψ_m}. It is interesting that this assumption is a sufficient condition for the measurement transformation, $P_{\phi_1} \longrightarrow \sum_m |S'_{nm}|^2 P_{\psi_m}$. To see this, recall that $\phi_1 = \sum_m S'_{nm}\psi_m$ so that E_m results with relative frequency $|S'_{nm}|^2$. Thus, if the post-measurement subensemble associated with result E_m is assumed to have state vector ψ_m it follows that the total post-measurement ensemble has density operator $\hat{\rho} = \sum_m |S'_{nm}|^2 P_{\psi_m}$. In Heisenberg's interpretation, therefore, the measurement transformation, the basis of which we are seeking, is not itself fundamental. Thus the question we have posed may be reformulated: why should a quantum theory of measurement assume that, after an A-measurement, the subensemble which yielded a_k has state vector α_k, $A\alpha_k = a_k\alpha_k$? For a "Copenhagen" explanation, we turn to Heisenberg's more recent philosophical works.

A common method of elucidating a complex subject is by analogy to something familiar, provided the analogy is not too superficial or purely

poetic. Thus we have seen in previous sections that features of quantum theory can be explained by drawing parallels to both classical optics and statistical mechanics. In the case of optics, the possibility of confusion loomed large; but the framework of statistical mechanics provided an excellent analogue to that of quantum theory--up to a point. Heisenberg has expounded the Copenhagen version of measurement theory by appealing to the latter analogy. Because of its evident impact upon quantum philosophers and theorists, we now closely scrutinize his analysis.

With Heisenberg,[26] consider first from the standpoint of classical Gibbsian statistics a hot metal occasionally emitting a thermal electron. Near this emitter is a photographic plate which registers all electrons emitted above some established threshold velocity. The temperature T of the metal is measured, and its thermodynamic state is represented mechanically by the canonical ensemble, i.e., $\rho(q,p) \propto \exp(-H(q,p)/kT)$, where H is the Hamiltonian of the metal. Now, as time passes, $\rho(q,p)$ develops in accordance with Liouville's equation; in particular, if the composite system of metal plus plate is considered, it is in principle possible to compute the probability that a given number of electrons have been detected by a certain time. Now, as noted in the preceding section on ensembles, it is possible to regard this use of the canonical ensemble as an expression of mechanical ignorance. Heisenberg clearly takes this position when he says that if an "observer is present, he will suddenly register the fact that the plate is blackened. The transition from the possible to the actual is thereby completed as far as he is concerned; he correspondingly alters the mathematical representation discontinuously, and the new ensemble contains only the blackened photographic plate...We see from this that the characterization of a system by an ensemble not only specifies the properties of this system, but also contains information

about the extent of the observer's knowledge of the system."[27]

To complete this classical analogue to the Copenhagen version of quantum measurement, it is necessary to provide a counterpart to complementarity. Following an idea of Bohr,[28] this may be done by recalling from statistical thermodynamics that a closed system is properly represented by a microcanonical ensemble, $\rho(q,p) \propto \delta(H(q,p)-E_0)$, whereas an open system (in thermal equilibrium with a heat reservoir) requires a canonical ensemble, $\rho(q,p) \propto \exp(-H(q,p)/kT)$. In the former case, the energy is fixed but the temperature is not determined; to measure the temperature, the system must be "opened" and put in thermal equilibrium with a thermometer. But when that is done, the energy fluctuates in accordance with the canonical distribution. Thus a macroscopic description involving the concept temperature is more or less "complementary" to a precise micromechanical description in which temperature is undetermined.

Obviously applying this "classical complementarity" to the hot metal and photographic plate, Heisenberg reasons that complete knowledge of the microstate of a closed metal-plus-plate system would permit exact rather than just probabilistic predictions concerning the blackening sequence, but "the statement of the temperature would then have been completely meaningless".[26] On the other hand, if that composite system is opened to its environment (called by Heisenberg "the external world"), then temperature supposedly becomes meaningful but precise knowledge of the microstate no longer eliminates probabilities, exact prediction being precluded because "we do not know every detail of the external world".[27]

The Copenhagen interpretation is essentially an attempt to provide exact quantal analogues for the concepts of statistical thermodynamics, provided the latter are understood in ways just explained.* A logical

*It should be noted that several of the above statements from classical statistics as interpreted by Heisenberg are familiar but not universally agreed upon by theoretical physicists.

first step would be to determine (1) what in statistical mechanics corresponds to the density operator, and (2) to what extent the analogy is correct. As we have already pointed out, as a mathematical object characterizing statistics of measurement results for an ensemble, the density operator plays the same role as Gibbs' density-of-phase; moreover, a quantum ensemble having a state vector (i.e., $\rho = P_\psi$) is analogous--so far as homogeneity is concerned--to a classical ensemble of systems all in the same mechanical state. But we have also observed that state vectors differ from classical microstates in ways (to be recalled as needed) which render this last analogy imperfect. Copenhagen theorists tend to ignore these infelicities; thus Heisenberg carries the ignorance interpretation of classical mixed ensembles over to quantum theory when he explains that "the probability function combines objective and subjective elements"[29] with this exception: "In ideal cases, the subjective element in the probability function may be practically negligible as compared with the objective one. The physicists then speak of a 'pure case'."[29] We have already called attention in section 5 to the inconsistency of this viewpoint. To regard a quantal mixture as expressing subjective ignorance of actual objective pure states is, in view of the essential latency of quantum observables, physically meaningless. Any attempt to assign pure states to individual elements of a mixed ensemble encounters hopeless ambiguity[4] primarily because of a deep logical fissure in the analogy to classical statistics, viz., the circumstance that in quantum theory homogeneity does not eliminate dispersion. The Copenhagen interpretation therefore pushes the analogy between density operator and density-of-phase beyond its proper bounds.

A second quantal analogue to statistical mechanics is based on the effects of interaction. We have systematically contrasted the dynamics of classical and quantal interactions elsewhere[4]; superficially the analogy

seems a good one, for in both theories an initially pure ensemble evolves into a mixture upon interaction with a mixed ensemble. In accordance with the ignorance interpretation of ensembles, Heisenberg therefore asserts that a system open to the "external world" must be described by a mixed ensemble, "since we do not know the details of the 'external world system'".[27] This reasoning is correct in classical physics but fallacious in quantum theory. Indeed, in the latter case, even if the "details were known" so that no "subjective" element entered the description of the "external world", i.e., even if the "external world" were in an objective, pure state, still the initially pure open system would evolve into a mixture![4]* Once again, quantum theory proves incompatible with the ignorance interpretation of ensembles.

In any case, only a closed**system can be dynamically characterized by a state vector; thus just as temperature was declared "meaningless" for a closed classical system, so apparently are all physical quantities for a closed quantum system. As Heisenberg puts it, although state vectors are objective, they are "abstract and incomprehensible", and "do not refer to real space or to a real property".[27] To make an actual measurement, system-plus-apparatus must be open, for "connection with the external world is one of the necessary conditions for the measuring apparatus to perform its function".[30] It follows of course that system-plus-apparatus can only be an element of a mixed ensemble; and for Heisenberg this automatically entails "statements about the observer's knowledge. If the observer later registers a certain behavior of the measuring apparatus as actual, he

*This is essentially why in section 6 the use of initially pure ensembles was sufficient to illustrate the basic features of quantum measurement theory.

**By closed we mean a system not interacting with its environment; i.e., the Hamiltonian for system plus environment has no interaction term.

thereby alters the mathematical representation discontinuously, because a certain one among the various possibilities has proved to be the real one."[30]

In mathematical terms, the Copenhagen description of an A-measurement on a system S from a pure ensemble therefore runs as follows. S is initially closed and in the state $\psi = \sum_k \langle \alpha_k, \psi \rangle \alpha_k$, which represents objective tendencies toward the possible A-values $\{a_k\}$ with respective probabilities $\{|\langle \alpha_k, \psi \rangle|^2\}$. However, since S is isolated, no A-measurement can be performed; for measurement requires interaction with surroundings. Now, an open system must be described by a density operator, i.e., by a whole ensemble of systems in various states. When S is "opened" for an A-measurement, its proper representative is therefore such an ensemble; but this introduces a subjective element, viz., ignorance as to which of the A-values actually obtains. The density operator after the measurement interaction (but before the actual A-value is observed) is accordingly $\hat{\rho} = \sum_k |\langle \alpha_k, \psi \rangle|^2 P_{\alpha_k}$, since a system in eigenstate α_k is certain to have the A-value a_k, and $|\langle \alpha_k, \psi \rangle|^2$ is just the probability originally associated with that value. Finally, observation of the actual A-value eradicates the ignorance; and, if a_k is the result, the state α_k is assigned to the system. The overall effect of an A-measurement upon the state $\psi = \sum_k \langle \alpha_k, \psi \rangle \alpha_k$ was therefore contraction of ψ to one term α_k, an act called by Copenhagen theorists "reduction of the wave packet".

There is the official answer to our question as to why derivation of the measurement transformation is popularly adopted as the goal of quantum measurement theory. That transformation, originally formulated by von Neumann, is indeed the correct mathematization of the Copenhagen philosophy of measurement. If the latter were espoused, the postulates of section 2 would have to be augmented by some statement connecting the measurement

concept to wave packet reduction in a definitive, analytic sense. However, since the foregoing dissection of the Copenhagen interpretation has revealed its foundation to be a set of overextended analogies to a highly subjective version of classical statistics, no supplementation of that kind seems at all justifiable.

9. Quantum Jumps

For historical reasons, the Copenhagen interpretation is effectively the orthodoxy of modern physics. Of course, the majority of physicists are not engaged in research so fundamental as that which originally motivated the Copenhagen philosophizing, and they are therefore not greatly interested in it. Nevertheless, when confronted with a basic theoretical dilemma, such as the quantum theory of measurement, most quantum theorists repeat at least some of the Copenhagen pronouncements, if only because of the suggestive phrases of workaday jargon. Perhaps the most common error, as we have already mentioned, is taking seriously expressions like "an electron in the state ψ", "the probability of finding an electron in the state ψ_n", and "the probability of a transition from state ψ_n to state ψ_m", which, although of practical, heuristic value, have picturesque connotations of describing a semiclassical microcosm alien to the spirit of quantum theory. In fact, as we demonstrated in the preceding section, the mathematical structure of quantum theory does not lend any support to a literal interpretation of such phrases.

Nevertheless, it is quite natural that most papers on the theory of measurement introduce the problem at least implicitly in the Copenhagen language. Hence, from our point of view, they are philosophically crippled at the outset and stand little chance of illuminating the measurement concept. The detailed criticism given in section 8 of Heisenberg's measurement

theory in particular serves therefore to expose generally the inadequacies of many discussions[31] on quantum measurement.

However, we cannot yet completely dismiss the mathematical condition placed upon measurement in the Copenhagen interpretation, provided different, and cogent, reasons can be given in its defense. A promising source appears at first to be one of Copenhagen's critics, Blochintsev,[32] whose typical Russian approach to quantum theory seems to have much in common with the early sections of the present work. In particular, he rejects the ignorance interpretation of quantum ensembles, prefering to regard membership in an ensemble as "objective". In an extraordinary rebuttal for someone who himself thoroughly dismantled the classical world view, Heisenberg has described Blochintsev's "objective" version of quantum ensembles as "taking us far--perhaps too far--from materialistic ontology",[33] because classically ensembles represented ignorance! At any rate, Blochintsev's objectivity is really political rather than mathematical, and the hypocrisy of his dialectical materialism finally emerges when he describes the measurement process as effecting precisely the state changes that Copenhagen decrees.[34] Unfortunately, no reasons are given; it is as though we simply need an additional axiom to characterize fully the idea of measurement:

P4: (Projection Postulate) If A-measurements are performed on an ensemble, the post-measurement subensemble consisting of those systems which yielded a_{kk} has density operator P_{kk}, $A\alpha_{kk} = a_{kk}\alpha_{kk}$.

P4 has been stated in the strongest form that could possibly make sense; it is sometimes carelessly expressed in terms of "sudden, acausal quantum jumps" or "superluminary wave packet contractions" having some vague relevance to the cognizance of events by observers, all of which is nonsense. Measurements themselves do not occur instantaneously, although

their results are associated with the moment that the apparatus couples to the system. P4 only suggests that when the interaction ceases, a certain selection of subensembles would always be possible. Together with the other axioms, it implies the measurement transformation $P_\psi \rightarrow \hat{\rho} = \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}$, as has already been proved; but we have found no reason to adopt either that transformation or P4 as a necessary property of measurement.

Occasionally, the classical proposition that an immediate repetition of an α-measurement which yielded a_R must again yield a_R is invoked in behalf of P4.[35] The only evidence for this assertion seems to be common sense (intuition drawn from classical physics). Fairly typical of this approach is a gedankenexperiment discussed by Heitler.[36] The arrangement is that of the electron diffraction experiment in section 7, except that a second flourescent screen is placed immediately behind the first one, now assumed to be quite thin. Suppose a glowing dot appears on the first screen, indicating an electron position measurement. Shortly a dot will burst forth on the second screen, but where? Heitler says it is "absurd" to believe the second dot could occur anyplace but directly behind the first one; hence "it follows that through the appearance of the electron on the first screen the probability distribution for the position must have changed and <u>contracted into one of certainty</u>".[36]

Superficially, this appears to be a reasonable argument in favor of quantum jumps, or of P4. Closer examination reveals, however, that even if such experiments were performed and the two dots were always together, the explanation would not fall beyond the scope of P1-P3; i.e., P4 would be unnecessary! To see this, the electron, the first flourescent screen, and the electromagnetic field must be considered as a single, composite quantum system. Initially, the electron is in a "traveling wave packet" state,

the atoms of the screen are in their ground states, and the electromagnetic field is in its vacuum state. After the interaction associated with the first position measurement, the state of this composite system is no longer so simply described; but we can determine, using only P1-P3, the probability P_{12} that the radiation field would appear as a glowing dot at some point x_1 on the first screen and that a second electron position measurement would yield a result x_2. To explain the effect used by Heitler to defend P4, all that would be needed is to show that P_{12} is vanishingly small unless $x_1 \sim x_2$, an analytic property completely derivable from P1-P3. This is reminiscent of the correct quantal explanation[37] of cloud chamber tracks, as opposed to the fanciful quantum jump version which interprets the tracks as sequences of position measurements, each collapsing the spreading wave packet to a point.

The ultimate appeal of this immediate remeasurement doctrine is to classical intuition; hence, in view of the essential latency of quantum observables and the concomitant nonpicturability of quantum systems, there is no a priori reason to believe that anything at all can be asserted about immediate remeasurement. However, the possibility is open that something like P4 might be <u>derivable</u> from P1-P3, since the latter have already endowed the construct measurement with several properties; and indeed, as we shall see below, such a theorem can <u>almost</u> be proved.

Since we are interested in the results of successive a-measurements, it would be convenient to have an auxiliary observable v_a whose operational definition involves such a measurement sequence. If a were a classical space coordinate, the velocity would be a suitable v_a; or, in general, for any classical a the time derivative $\frac{da}{dt}$ is evaluated by successive a-measurements. We therefore consider the quantal analogue of $\frac{da}{dt}$, attempting to make inferences from its operator and eigenvalues.

To find the operator $\frac{dA}{dt}$ which corresponds to $\frac{da}{dt}$ (if any operator does), it is sufficient to require that $\frac{dA}{dt}$ satisfy the classically-inspired statistical equation $\frac{d}{dt}m(A) = m\left(\frac{dA}{dt}\right)$. It is then a standard quantum theorem,[38] based on P3 and Th1, that $\frac{dA}{dt} = \frac{1}{i\hbar}[A,H] + \frac{\partial A}{\partial t}$, where $[A,H] \equiv AH - HA$ and H is the Hamiltonian of the system (cf. fn., p. 7). By definition, a is conserved relative to H if $\frac{d}{dt}m(A) = 0$, for every admissible functional m; hence a is conserved if $\frac{dA}{dt} = 0$.

Against this background, consider successive measurements of a conserved observable a performed at t_1 and $t_1 + \Delta t_1$ upon a system from an initial ensemble described by $\rho(t_1)$. Since a is conserved, $\frac{dA}{dt} = 0$; therefore Th4 implies that measurement of $\frac{da}{dt}$ must yield zero. So far, the assumed operational definition of $\frac{da}{dt}$ has been drawn from the classical case; thus, $\frac{dA}{dt} = 0$ must mean that two successive a-measurements at t_1 and $t_1 + \Delta t_1$ will both yield the same result in the limit $\Delta t_1 \to 0$, in accordance with the classical definition, $\frac{da}{dt} \equiv \lim_{\Delta t_1 \to 0} \frac{a(t_1 + \Delta t_1) - a(t_1)}{\Delta t_1}$. Have we therefore derived the projection "postulate" (for conserved observables) from the basic quantal axioms? If so, an interesting corollary would be that P4 is always false for systems in which the observable of interest B is such that $\frac{dB}{dt}$ has no zero eigenvalues.

However, neither of these conclusions is reliable. To see why, consider carefully the reasoning which leads to $\frac{d}{dt}m(A) = m\left(\frac{dA}{dt}\right)$, the defining equation for $\frac{dA}{dt}$ in quantum theory. That equation follows from the purely classical picture of a function (classical possessed observable) evolving in time; the "measurements" of $a(t_1)$ and $a(t_1 + \Delta t_1)$ are naturally assumed not to disturb the causal development of $a(t)$, i.e., to leave the classical state unchanged. The foregoing quantum theoretical

interpretation of $\frac{da}{dt}$ was therefore implicitly based on the notion that successive a-measurements could be performed on the systems of an ensemble without in any way disturbing its causal evolution, i.e., without significant interaction. As might be expected, that assumption leads to a contradiction, as follows. Suppose $\rho(t_1) = P_{4(t_1)}$ and a-measurements are performed at t_1 (a conserved). These two propositions follow: (1) Since the a-measurements were assumed not to alter the normal temporal development of the ensemble, $\rho(t_1+\Delta t_1) = P_{4(t_1+\Delta t_1)}$, i.e., the ensemble must remain <u>homogeneous</u>. (2) By the "theorem" just proved, any system which yielded a_k in the a-measurement at t_1 must do so again at $t_1+\Delta t_1$, $\Delta t_1 \to 0$; therefore, immediately after the first measurement, the ensemble can be subdivided into subensembles, each having the property that an a-measurement will yield a given result with <u>certainty</u>. These subensembles are then distinct and homogeneous; hence the total ensemble is <u>inhomogeneous</u>.

The statements (1) and (2) obviously contradict each other; the above "derivation" of P4 for conserved observables is therefore illegitimate. However, it is not entirely worthless; for it essentially affirms that the "observable" $\frac{da}{dt}$ as conceived by classical intuition is <u>unobservable</u> in quantum theory. And this perhaps indicates that quantum measurements on an ensemble will in general cause its density operator to deviate from undisturbed temporal development. However, instead of attempting to draw basic conclusions about the quantum measurement concept from the foregoing mathematical skirmish, it seems preferable to regard it simply as an illustration of the difficulties inherent in the application of classical pictures to quantum problems.

Lest the foregoing remarks be interpreted as an assertion that velocity, for example, is unobservable, we hasten to add a few remarks clarifying

the meaning of that construct in quantum theory. As just established, velocity cannot have the same empirical meaning both classically and quantally, because of fundamental differences in the two disciplines. On the other hand, neither should the construct be assigned a radically different meaning for each of the two. The situation is rather like an old problem in geometry, viz., defining parallelism. In the Euclidean case, the concept parallel is well entrenched, being endowed with a number of seemingly immutable qualities. Generalized Riemannian geometries, among which Euclid's is a special case, naturally should incorporate somehow a generalized parallelism; but this is impossible if all of the properties traditionally associated with that concept are insisted upon. Accordingly, only a few of them are used in the definition of equipollence, which is applicable to all Reimannian geometries and which, in the Euclidean system, reduces to parallelism.

Similarly, in quantum theory, velocity is introduced by seizing upon a single property of the original concept: $\frac{d}{dt}m(X) = m\left(\frac{dX}{dt}\right)$ where X is the position operator and $\frac{dX}{dt}$, the velocity operator. Admittedly, the relation between classical and quantum is not strictly the same as that between Euclidean and Reimannian; the former connection involves an approximation, the latter does not. However, just as equipollence must become parallelism in the Euclidean case, $\frac{dX}{dt}$, to deserve the name velocity, must satisfy appropriate equations in the classical limit. The Ehrenfest[39] theorems establish that it does. Thus our statement that the classically defined observable $\frac{dx}{dt}$ is quantally unobservable does not mean that velocity is unobservable; velocity is the quantum observable whose operator is $\frac{dX}{dt}$ and which is equivalent to the classical velocity in every way only in the classical limit.

At any rate, our search for a formal derivation of P4 was unsuccessful;

hence we still have no persuasive argument in behalf of its adoption as a basic postulate.

10. An Untenable Consequence of the Standard Theory

In its crude form based on assignment of state vectors to single systems, the projection postulate is easily used to "prove" that simultaneous measurement of noncommuting observables is impossible. All that is required is the observation that the post-measurement state, by virtue of wave packet reduction, would have to be simultaneously an eigenvector of two different operators. Unless the latter commute, such an eigenvector is a rarity; in fact usually none exists at all. Now, we hold that any theorem purporting to prove the impossibility of simultaneous measurement of two observables is necessarily founded on a false hypothesis. This follows from the fact that it is possible to construct within the quantal framework given by P1-P3 legitimate models* of simultaneous measurement schemes for noncommuting observables. Thus from a physical point of view, to say that the statement "$\mathcal{A}$ and $\mathcal{B}$ cannot be measured simultaneously" is an analytic truth simply condemns the axiom set from which it was derived. We have then essentially a reductio ad absurdum argument against the fanciful version of wave packet reduction, which has, however, already been rejected earlier on other grounds.

This raises the question as to whether the weaker (but strongest admissible) form of the projection postulate, stated in section 9 as P4, is also subject to such a critique; the answer is that P4, as an isolated postulate, is not directly assailable along these lines because of its careful association of eigenvectors with ensembles rather than individual systems.

---

*Such counterexamples to the widely believed principle of incompatibility of noncommuting observables are discussed in detail in Ref. 2.

Nevertheless, it turns out that the measurement transformation, a consequence of P4, does become untenable when confronted with basic quantal theorems; for like naive wave packet reduction, it implies that two observables $\mathcal{A}$ and $\mathcal{B}$ are simultaneously measurable only if $[A, B] = 0$.

To prove this, recall first that according to von Neumann's measurement transformation, an $\mathcal{A}$-measurement on a pure ensemble converts the density operator from P_ψ to the mixture $\sum_R |\langle \alpha_R, \psi\rangle|^2 P_{\alpha_R}$. Similarly, a $\mathcal{B}$-measurement would induce the change $P_\psi \longrightarrow \sum_\ell |\langle \beta_\ell, \psi\rangle|^2 P_{\beta_\ell}$. If this transformation is a universal property of measurement, a simultaneous measurement of $\mathcal{A}$ and $\mathcal{B}$ must therefore be described by

$$P_\psi \to \hat{\rho} = \sum_R |\langle \alpha_R, \psi\rangle|^2 P_{\alpha_R} = \sum_\ell |\langle \beta_\ell, \psi\rangle|^2 P_{\beta_\ell}.$$

According to the probability rules of quantum theory, $W_{\mathcal{B}}(b_\ell; \hat{\rho}) = Tr(\hat{\rho} P_{\beta_\ell})$, which is independent of the particular representation of $\hat{\rho}$. Hence,

$$W_{\mathcal{B}}(b_\ell; \hat{\rho}) = Tr(\hat{\rho} P_{\beta_\ell}) = \sum_R |\langle \alpha_R, \psi\rangle|^2 Tr(P_{\alpha_R} P_{\beta_\ell}) = \sum_R |\langle \alpha_R, \psi\rangle|^2 |\langle \beta_\ell, \alpha_R\rangle|^2,$$

but also $W_{\mathcal{B}}(b_\ell; \hat{\rho}) = Tr(\hat{\rho} P_{\beta_\ell}) = |\langle \beta_\ell, \psi\rangle|^2$; thus we have

$$|\langle \beta_\ell, \psi\rangle|^2 = \left|\sum_R \langle \beta_\ell, \alpha_R\rangle \langle \alpha_R, \psi\rangle\right|^2 = \sum_R |\langle \alpha_R, \psi\rangle|^2 |\langle \beta_\ell, \alpha_R\rangle|^2.$$

Clearly, this does not hold for every ψ, $\{\alpha_R\}$, and $\{\beta_\ell\}$. To find conditions under which it is correct, note that

$$\left|\sum_R \langle \beta_\ell, \alpha_R\rangle \langle \alpha_R, \psi\rangle\right|^2 = \left(\sum_R \langle \beta_\ell, \alpha_R\rangle \langle \alpha_R, \psi\rangle\right)^* \left(\sum_n \langle \beta_\ell, \alpha_n\rangle \langle \alpha_n, \psi\rangle\right)$$
$$= \sum_R |\langle \alpha_R, \psi\rangle|^2 |\langle \beta_\ell, \alpha_R\rangle|^2 + \sum_{n,R,n\neq R} \langle \alpha_R, \beta_\ell\rangle \langle \beta_\ell, \alpha_n\rangle \langle \psi, \alpha_R\rangle \langle \alpha_n, \psi\rangle.$$

The measurement transformation is therefore applicable only to simultaneous $\mathcal{A}, \mathcal{B}$ measurements such that

$$(1) \quad \sum_{n\neq R} \langle \alpha_R, \beta_\ell\rangle \langle \beta_\ell, \alpha_n\rangle \langle \psi, \alpha_R\rangle \langle \alpha_n, \psi\rangle = 0,$$

which means that the measurement transformation can describe simultaneous measurements only for some pairs of observables, viz., those whose eigenvectors satisfy equation (1) for all ψ.

To find the relation between the sets of eigenvectors $\{\alpha_R\}$ and $\{\beta_\ell\}$,

we consider special ψ's:

Let $\psi = \frac{1}{\sqrt{2}}(\alpha_N + \alpha_K)$, $N \neq K$. Substitution into (1) gives

$$\sum_{n \neq R} \langle \alpha_R, \beta_\ell \rangle \langle \beta_\ell, \alpha_n \rangle \left(\tfrac{1}{2}\right)(\delta_{RN} + \delta_{RK})(\delta_{nN} + \delta_{nK})$$
$$= \tfrac{1}{2}\left(\langle \alpha_N, \beta_\ell \rangle \langle \beta_\ell, \alpha_K \rangle + \langle \alpha_K, \beta_\ell \rangle \langle \beta_\ell, \alpha_N \rangle\right)$$
$$= \mathrm{Re}\left(\langle \alpha_N, \beta_\ell \rangle \langle \beta_\ell, \alpha_K \rangle\right) = 0.$$

Similarly, let $\psi = \frac{1}{\sqrt{2}}(\alpha_N + i\alpha_K)$, $N \neq K$, to get

$$\sum_{n \neq R} \langle \alpha_R, \beta_\ell \rangle \langle \beta_\ell, \alpha_n \rangle \left(\tfrac{1}{2}\right)(\delta_{RN} - i\delta_{RK})(\delta_{nN} + i\delta_{nK})$$
$$= \tfrac{i}{2}\left(\langle \alpha_N, \beta_\ell \rangle \langle \beta_\ell, \alpha_K \rangle - \langle \alpha_K, \beta_\ell \rangle \langle \beta_\ell, \alpha_N \rangle\right)$$
$$= -\mathrm{Im}\left(\langle \alpha_N, \beta_\ell \rangle \langle \beta_\ell, \alpha_K \rangle\right) = 0.$$

Hence, equation (1) may be replaced by the simpler restriction

(2) $\langle \alpha_n, \beta_\ell \rangle \langle \beta_\ell, \alpha_m \rangle = 0$, for every ℓ, n, m, $n \neq m$,

Since $\{\alpha_R\}$ and $\{\beta_\ell\}$ are <u>eigenvector</u> sets, neither contains the null vector. Thus, any α_M has the property that $\langle \beta_\ell, \alpha_M \rangle \neq 0$ for <u>some</u> value of ℓ , say $\ell = L$. Equation (2) then implies that $\langle \alpha_n, \beta_L \rangle = 0$, $n \neq M$, i.e., β_L is orthogonal to every element of $\{\alpha_R\}$ except one, α_M . But $\{\alpha_R\}$ is a complete set; therefore α_M and β_L must be equal up to a phase factor (i.e., belong to the same ray in Hilbert space). The same argument applies to all values of L. Hence, each element of complete set $\{\beta_\ell\}$ is an element of complete set $\{\alpha_R\}$; the eigenvector sets for observables $\mathcal{A}$ and $\mathcal{B}$ are identical (except for unimportant phases). It follows that the operators A, B corresponding to simultaneously measurable observables $\mathcal{A}, \mathcal{B}$ <u>must commute</u>!

We have therefore proved that the measurement transformation $P_\psi \rightarrow \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}$ <u>forbids</u> the simultaneous measurement of noncommuting observables.

Yet it is just this transformation which is widely accepted as a universal characteristic of measurement. Most presentations of the quantum theory of measurement, including von Neumann's and London and Bauer's

classic treatments, adopt it as a goal. A derivation of it typically counts not only as a general explanation of measurement but also as a demonstration of the internal consistency of quantum theory. We now see that this point of view must be rejected, for the transformation entails an absurd corollary, viz., that simultaneous measurements of noncommuting observables are impossible. Hence the transformation $P_\psi \to \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}$ cannot be upheld as a defining attribute of the quantum emasurement process.

11. Inadequacy of the Correlation Assumption

An interesting consequence of the measurement transformation is the stringent condition it places on A-measurements performed upon a pure ensemble with state vector α_n i.e., measurements certain to yield a_n. According to that transformation, such an ensemble is left unaltered by measurement: if $\rho = P_{\alpha_n}$, $\hat{\rho} = \sum_R P_{\alpha_R} P_{\alpha_n} P_{\alpha_R} = \sum_R P_{\alpha_R} \delta_{Rn} = P_{\alpha_n}$. Consider again the general schema of section 6, where measurement was depicted as an interaction between system S and apparatus M. Let the states of S and M be α_R and χ_0 at the beginning of the A-measurement interaction. This dynamical process is represented generally by $T_A(\alpha_R \otimes \chi_0) = \theta(\alpha_R, \chi_0)$; but, since $\theta(\alpha_R, \chi_0)$ must be the Rth eigenvector of $A \otimes 1$, it follows that $\theta(\alpha_R, \chi_0) = \alpha_R \otimes \chi(\alpha_R, \chi_0)$. We have therefore proved the following theorem: the correlation assumption, $T_A(\alpha_R \otimes \chi_0) = \alpha_R \otimes \chi(\alpha_R, \chi_0)$, is a necessary condition for the measurement transformation.

If $\psi = \sum_R \langle \alpha_R, \psi \rangle \alpha_R$ is the initial state of S, the correlation assumption becomes

$$T_A(\psi \otimes \chi_0) = \sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \chi(\alpha_R, \chi_0),$$

from which the post-measurement density operator $\hat{\rho}$ for the S-ensemble is easily calculated. Let $\{\varphi_n\}$ be a complete orthonormal set in $\mathcal{H}_2$.

The desired density operator is then given by

$$\hat{\rho}_1 = \mathrm{Tr}_2 P_{\Sigma_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \chi_R}, \quad \chi_R \equiv \chi(\alpha_R, \chi_0):$$

$$\hat{\rho}_1 = \sum_{n,R,\ell} \langle \vartheta_n | \langle \alpha_R, \psi \rangle \alpha_R \otimes \chi_R \rangle \langle \langle \alpha_\ell, \psi \rangle \alpha_\ell \otimes \chi_\ell | \vartheta_n \rangle$$

$$= \sum_{n,R,\ell} \langle \alpha_R, \psi \rangle \langle \psi, \alpha_\ell \rangle \, |\alpha_R\rangle\langle\alpha_\ell| \, \langle \vartheta_n, \chi_R \rangle \langle \chi_\ell, \vartheta_n \rangle$$

$$= \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R} \Big(\sum_n |\langle \vartheta_n, \chi(\alpha_R, \chi_0) \rangle|^2 \Big)$$

$$+ \sum_{R,\ell, R \neq \ell} \langle \alpha_R, \psi \rangle \langle \psi, \alpha_\ell \rangle \, |\alpha_R\rangle\langle\alpha_\ell| \, \langle \chi(\alpha_R, \chi_0), \chi(\alpha_\ell, \chi_0) \rangle.$$

The factor in parentheses is unity since $\chi(\alpha_R, \chi_0)$ is normalized; the correlation assumption would therefore have the measurement transformation as a consequence if the $R \neq \ell$ summation vanished. To eliminate the cross terms it is adequate to require orthogonality of the M-states $\chi(\alpha_R, \chi_0)$ which are correlated to different S-states α_R. The theory of measurement outlined in section 6 employed the even stronger statement that $\{\chi(\alpha_R, \chi_0)\} = \{\gamma_R\}$, eigenvectors of an observable belonging to M, which is physically an essential requirement since without it M could not play the role of an a-meter. Thus the correlation assumption in the form $T_A(\alpha_R \otimes \chi_0) = \alpha_R \otimes \gamma_R, \{\gamma_R\}$ an orthogonal set, is sufficient to derive the measurement transformation.

Combining this with the preceding theorem, we now see that the correlation assumption is necessary and sufficient for, hence equivalent to, the measurement transformation. Our recurrent question concerning the universal validity of the measurement transformation may therefore be rephrased as follows: does the correlation assumption describe a necessary attribute of quantum measurement? This version will shortly become the center of attention.

Incidentally, this effective equivalence of $T_A(\alpha_R \otimes \chi_0) = \alpha_R \otimes \gamma_R$ and $P_4 \to \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R}$ enables the deduction of P4 as a consequence of the measurement transformation, which seemed only plausible until now. (The converse was proved in section 8.) To prove this, note

that $T_A(\psi \otimes \chi_0) = \sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \gamma_R$ is supposed to describe the A-measurement procedure in the sense that a C-measurement on M is sufficient to predict with certainty what a concurrent A-measurement on S would have given. The certainty of the prediction is assured by the correlated form of $\sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \gamma_R$ which accords zero probability to pairs of measurement results (a_n, c_m), if $n \neq m$. This certain prediction is then called the result of an A-measurement at the instant S and M began to interact. Consider now the S-subensemble defined by the fact that A-measurements all "yielded" a_R in the sense just explained. It follows that after the measurement interaction such a subensemble would yield a_R with certainty if subjected to A-measurement, hence must be assigned density operator P_{a_R}. Within the standard framework of measurement theory as outlined in section 6, we have therefore established the equivalence of the projection postulate, the measurement transformation, and the correlation assumption.

As in the case of the measurement transformation, there are severe theoretical difficulties inherent in the correlation assumption. First, it defies generalization to simultaneous measurements of arbitrary observables, as we have shown elsewhere.[40] Second, there is an interesting theorem about the correlation assumption which proves that such correlation cannot be a universal attribute of measurement because it requires an extremely rare type of physical interaction. A special case was first discovered by Wigner;[41] later Araki and Yanase[42] gave a general proof, including even the possibility of degenerate eigenvectors in the correlation assumption. However, the essential structure of the theorem may be understood without the latter refinement.

The major point is that virtually every physical interaction displays some kind of symmetry, with a conservation law as its dynamical consequence.

In particular, it would be completely unreasonable to claim that no measurement interaction could have a conserved observable. Accordingly, let $L = L_1 \otimes 1 + 1 \otimes L_2$ be conserved relative to the correlation process, $T_A(\alpha_n \otimes \chi_0) = \alpha_n \otimes \gamma_n$. If L is not explicitly time dependent, this means that $[T_A, L] = 0$. The Wigner-Araki-Yanase theorem may be stated as follows (omitting degeneracies):

WAY Th: If (1) $T_A(\alpha_n \otimes \chi_0) = \alpha_n \otimes \gamma_n$, $\langle \gamma_k, \gamma_m \rangle = \delta_{km}$, and (2) there exists $L = L_1 \otimes 1 + 1 \otimes L_2$ such that $[T_A, L] = 0$ then $[L_1, A] = 0$.

The system observable A, which is being measured, must commute with every conserved observable! For later reference, the essentials of the proof will now be reviewed; however, to clarify the range of applicability of the theorem, we shall take a slightly generalized form of the correlation assumption, viz., $T_A(\alpha_n \otimes \chi_0) = \beta_n \otimes \gamma_n$, which opens the possibility that measurement might leave a system initially in eigenstate α_n in a different state β_n.

Outline of the proof:

$$(1)\ \langle \alpha_n \otimes \chi_0 | L | \alpha_m \otimes \chi_0 \rangle = \langle T_A(\alpha_n \otimes \chi_0), T_A L (\alpha_m \otimes \chi_0) \rangle$$
$$= \langle T_A(\alpha_n \otimes \chi_0), L T_A(\alpha_m \otimes \chi_0) \rangle = \langle \beta_n \otimes \gamma_n | L | \beta_m \otimes \gamma_m \rangle$$
$$= \langle \beta_n, L_1 \beta_m \rangle \langle \gamma_n, \gamma_m \rangle + \langle \beta_n, \beta_m \rangle \langle \gamma_n, L_2 \gamma_m \rangle.$$

$$(2)\ \langle \alpha_n \otimes \chi_0 | L | \alpha_m \otimes \chi_0 \rangle = \langle \alpha_n, L_1 \alpha_m \rangle \langle \chi_0, \chi_0 \rangle + \langle \alpha_n, \alpha_m \rangle \langle \chi_0, L_2 \chi_0 \rangle.$$

From (1) and (2) is obtained an expression for the matrix elements of in the A-representation:

$$(3)\ \langle \alpha_n, L_1 \alpha_m \rangle = \langle \beta_n, L_1 \beta_m \rangle \delta_{nm} + \langle \beta_n, \beta_m \rangle \langle \gamma_n, L_2 \gamma_m \rangle - \delta_{nm} \langle \chi_0, L_2 \chi_0 \rangle.$$

Now, to prove that $[L_1, A] = 0$, it is sufficient to demonstrate that L_1 commutes with the spectral projectors of A, i.e., with all the P_{α_k}.

Thus we must show that $\langle \alpha_n, P_{\alpha_k} L_1 \alpha_m \rangle$ and $\langle \alpha_n, L_1 P_{\alpha_k} \alpha_m \rangle$ are equal, for every k, n, m.

(4) $\langle \alpha_n, P_{\alpha_k} L_1 \alpha_m \rangle = \delta_{kn} \langle \alpha_n, L_1 \alpha_m \rangle$.

(5) $\langle \alpha_n, L_1 P_{\alpha_k} \alpha_m \rangle = \delta_{km} \langle \alpha_n, L_1 \alpha_m \rangle$.

Subtract (5) from (4): $\langle \alpha_n, [P_{\alpha_k}, L_1] \alpha_m \rangle = \langle \alpha_n, L_1 \alpha_m \rangle (\delta_{kn} - \delta_{km})$.

Clearly if $n = m$, $\langle \alpha_n, [P_{\alpha_k}, L_1] \alpha_m \rangle = 0$. For $n \neq m$, note from (3) that

$$\langle \alpha_n, L_1 \alpha_m \rangle = \langle \beta_n, \beta_m \rangle \langle \gamma_n, L_2 \gamma_m \rangle, \quad n \neq m.$$

In the original form of the correlation assumption, $\beta_n = \alpha_n$, hence $\langle \beta_n, \beta_m \rangle = 0$, $n \neq m$, and the theorem is proved.

For the generalized form, however, the theorem can be valid only if an additional restriction is made which leads to $\langle \beta_n, \beta_m \rangle \langle \gamma_n, L_2 \gamma_m \rangle = 0$, $n \neq m$. Orthogonality of the $\{\beta_n\}$ is obviously the most natural assumption to make, and that can be assured if $\{\beta_n\}$ is regarded as the eigenvector set for some observable.

How is the WAY theorem to be interpreted? What is its impact on the standard theory of measurement? Strangely enough, its authors have been remarkably conservative in their assessments of its significance by hastening to prove that, although the correlation assumption is rigorously untenable, it can nevertheless be approximately valid.[41,42] Thus Yanase[43] regards the theorem as a basis for classifying observables: (1) measurable (those which commute with all conserved quantities), (2) approximately measurable (those which do not), and (3) unmeasurable (those conjugate to an observable belonging to a superselection rule). In connection with the postulate P1 (cf. fn., p. 7), (3) may be expressed in another way: operators which do not commute with those belonging to superselection rules do not correspond to observables. This seems preferable to associating them with "unmeasurable observables".

Moreover, we cannot accept the premise which evidently underlies

Yanase's first and second classes, viz., the proposition that the correlation assumption is still the "correct" theoretical definition of measurement even though that assumption is demonstrably false in almost all cases. Rather than maintain the Copenhagen-inspired belief that the essence of all true measurement is embodied in the correlation assumption (or equivalently, the measurement transformation or projection postulate), why not accept the full logical force of the WAY theorem? That theorem simply proves that the standard theory of measurement, i.e., the one created by von Neumann and reported by London and Bauer and others, is wrong; the correlation assumption central to that theory does not define the term measurement in general. The theory probably does of course apply to some measurement processes, but it certainly does not apply to all. To contrast Yanase's interpretation of the WAY theorem with that given here, consider an operator A which does not commute with all the operators conserved relative to any interaction compatible with the correlation assumption. Yanase would say that the observable a is only approximately measurable in principle; we claim that all observables are measurable (exactly) in principle, and that no theorem could possibly refute this elementary connection between the constructs observable and measurement. What the WAY theorem does affirm is a conclusion which foregoing sections have strongly suggested: the correlation assumption and equivalent statements do not express a defining attribute of the term measurement.

Let us now briefly recapitulate the major developments and conclusions thus far. We are engaged of course in a study of the nature and meaning of the concepts measurement and preparation in quantum theory. Having observed that the von Neumann transformation $P_\psi \rightarrow \sum_r |\langle \alpha_r, \psi \rangle|^2 P_{\alpha_r}$ plays a central role in virtually all descriptions of the quantal measurement process, we have sought in vain to justify its adoption as a necessary

property of measurement. This search led to the influential Copenhagen interpretation of measurement, which was therefore subjected to intense scrutiny, but with the result that no palatable explanation was to be found there. A temporary shift in attention from the measurement transformation to the essentially equivalent projection "postulate" also bore no fruit. In section 10, however, we exposed a hidden consequence of the measurement transformation which entails a strong theoretical argument against its universal validity; thus a damaging lacuna was revealed in the standard form of quantum measurement theory in connection with its extensibility to simultaneous measurements. Finally, the present section introduced another equivalent formulation of our basic question in terms of the correlation assumption. The WAY theorem was then invoked to demonstrate the inadequacy of that assumption as a universal description of quantum measurement. We therefore conclude that the standard theory of measurement, which is an outgrowth of the Copenhagen interpretation, is not of general validity, i.e., it is not really the quantum theory of measurement at all but at best a description of a relatively small class of possible measurement processes.

12. Alternate Correlation Schemes

In one of the earliest treatments of quantum measurement theory, Pauli[44] recognized that not all measurement processes could be subsumed under the same correlation assumption, $T_A^{(1)}(\alpha_R \otimes \chi_0) = \alpha_R \otimes \beta_R$. Those that did fit this pattern were termed measurements of the first kind; their basic property was taken to be that the initial state of the system was also its post-measurement state. Pauli apparently reasoned that other measurements always involved a change in (pure) state from α_R to β_R, for he defined only this one additional class, measurements of the second

kind: $T_A^{(2)}(\alpha_{kR} \otimes \chi_0) = \beta_{kR} \otimes \gamma_{kR}$; $\{\beta_{kR}\}$ is taken to be the eigenvector set for some observable.

Supposedly the Franck-Hertz determination of atomic energy levels exemplifies the second kind of measurement; however, this is quite impossible unless that experiment is described in the subjective Copenhagen language complete with wave packet reduction.*

As proof of this minor point and as an important example of a real measurement process, the basic quantum mechanics of the Franck-Hertz energy measurements will now be reviewed. It is assumed that the projectile electron is initially in energy eigenstate χ_0 and the target atom in (unknown) energy eigenstate α_{kR}. The purpose of the measurement procedure is to find k. When worked out, the collision process has this dynamical representation: $T_A(\alpha_{kR} \otimes \chi_0) = \sum_\ell d_\ell^{(kR)} \beta_\ell \otimes \gamma_\ell$, where $\{\beta_\ell\}$ is just a different ordering of the energy eigenstates $\{\alpha_{kR}\}$, and $\{\gamma_\ell\}$ is the set of energy eigenstates of the scattered electron. (In this highly schematized discussion, no effort is made to distinguish sums and integrals, or discrete and continuous indices.) The sum of energy eigenvalues belonging to each correlated eigenvector pair $\beta_\ell \otimes \gamma_\ell$ is equal to that for the initial state $\alpha_{kR} \otimes \chi_0$; this is just energy conservation in a quantal context.**

Thus if the energy spectrum of the atom is known, and an energy measurement is performed on the scattered electron, it is possible to predict

---

*Pauli himself did accept wave packet reduction, hence was not forced to define his second kind of measurement solely by an evolution operator $T_A^{(2)}$. Nevertheless, the name Pauli is occasionally attached, as above, to $T_A^{(2)}$, although that operator is not to be found in his work. The present critique thus refers to $T_A^{(2)}$, but not to Pauli.

**Incidentally, all quantum conservation laws are similarly derivable. Occasionally one sees the erroneous contention[45] that quantum theory implies only that mean values are conserved, and thus needs an extra postulate to account for conservation in elementary scattering processes.

with certainty the potential result of a concurrent energy measurement on the atom. The initial state α_{kk} is then inferred from the conservation law, and its eigenvalue is declared the result of this energy measurement. As explained in detail in section 8, Copenhagen theorists would now say that the measurement of the scattered electron has revealed which of several possible "transitions" from state α_{kk} has actually occurred, and the wave packet $\sum_l d_l^{(kk)} \beta_l \otimes \gamma_l$ is accordingly reduced to $\beta_{kk} \otimes \gamma_{kk}$. Only with this unwarranted interpretation can the Franck-Hertz experiment be said to illustrate so-called measurements of the second kind. (Even then, we must write $T_A^{(2)}(\alpha_{kk} \otimes \chi_0) = \sum_l d_l^{(kk)} \beta_l \otimes \gamma_{kk} \rightarrow \beta_{kk} \otimes \gamma_{kk}$ where the arrow denotes a <u>mental</u> process.) At any rate, even if there are measurements which do satisfy the condition, $T_A^{(2)}(\alpha_{kk} \otimes \chi_0) = \beta_{kk} \otimes \gamma_{kk}$, $\{\beta_{kk}\}$ orthogonal, the WAY theorem, as we noted in the last section, is applicable to them. Hence, the so-called Pauli classifications are far from complete. In fact, as we have seen in the Franck-Hertz case, they do not even include measurements by inelastic scattering, perhaps the most common experiment in modern physics.

There is, however, a variant of the correlation assumption which the WAY theorem does not rule out. From the proof given in section 11, it is clear that if $\{\beta_{kk}\}$ is <u>not</u> an orthogonal set, nothing can be said against a modified correlation assumption of this form:

$T_A(\alpha_{kk} \otimes \chi_0) = \beta_{kk} \otimes \gamma_{kk}$, $\{\beta_{kk}\}$ not necessarily an orthogonal set. In fact, this version is central to the quantum measurement theory presented by Landau and Lifshitz;[46] but, like Pauli's correlation, it cannot be upheld as a <u>general</u> condition on measurement, inelastic scattering being once again a good counterexample.

The foregoing sequence of proposed correlation assumptions almost suggests that we are gradually approaching a "correct" one upon which a

comprehensive quantum theory of measurement will be founded. Thus Pauli's first kind was a special case of his second kind; and the latter, a special case of the Landau and Lifshitz correlation. We might therefore conjecture that the proper all embracing assumption should have some vague form like $T_A(\alpha_R \otimes \chi_0) = \Sigma_{R0}$, for example, which has the merit of including every correlation discussed in this section, including the Franck-Hertz case. Still another scheme has been proposed by Durand,[47] who almost succeeded in giving a general description of the measurement process by correlating A-eigenvectors to mean values of a quasi-classical observable associated with M; but that theory turned out to be applicable only to the so-called selective measurements (idealized analogues of the Stern-Gerlach experiment). However, we shall finally abandon this quest for an elusive ultimate correlation assumption, for it is actually quite pointless; as will be explained later, even if a sufficiently general, or vague, correlation were discovered, it could not possibly serve as a definition of the term measurement as used in the quantum axioms.

13. The Apparatus as a "Classical" System

When Schrödinger offered his "burleske"[48] quantum description of a cat in a box, he illustrated a point which many quantum theorists have taken seriously in connection with measurement theory. Schrödinger's cat is incarcerated in a chamber containing a few radioactive atoms and some equipment. The only interaction between cat and atoms occurs when an atom disintegrates, but that rare event will trigger some lethal machinery. A geiger counter responds to the decay by setting into operation a hammer which shatters a flask of cyanide. Thus the interaction correlates the possible states of the atoms with the "alive" and "dead" states of the cat.

Consider now a quantum theoretical description of the composite system

consisting initially of the cat and one unstable atom. The cat-observable of interest is the proposition, "it is alive", with two eigenvalues, 1 (yes) and 0 (no), to which belong eigenstates α and δ, respectively. For the atom, let ϕ denote its initial unstable state and θ its possible stable state. At first, the composite system is in the state $\psi(0)=\alpha\otimes\phi$; as time progresses, ψ develops into $\psi(t)=C_1(t)\alpha\otimes\phi+C_2(t)\delta\otimes\theta$, which indicates the correlation between the two systems.

Reluctance to accept this as an adequate description of what has happened in the box stems from the unfortunate literal interpretation of the phrase "a system in state ψ", which we have criticized repeatedly. Thus it is said[49] that in actuality the state of a cat is never a blurred superposition of "living" and "dead" eigenstates, but is at all times one or the other, though which one might be unknown. To express this ignorance (recall the Copenhagen interpretation), a mixture is required. Even though we contend that this demand arises from a misinterpretation of ψ, nevertheless it seems at first that this mysterious desire for a mixture to describe the cat is automatically fulfilled anyhow. Indeed, the mortality statistics for the cat alone are easily calculated and the density operator is, as a matter of fact, mixed:

$$\rho_{cat}(t)=|C_1(t)|^2P_\alpha+|C_2(t)|^2P_\delta.$$

Strangely enough, this does not satisfy the objectors; their demands are even stronger. Supposedly, it is an a priori truth that a cat-atom system should be described by the mixed, correlated density operator,

$$\rho(t)=|C_1(t)|^2P_{\alpha\otimes\phi}+|C_2(t)|^2P_{\delta\otimes\theta},$$

which would refer of course to an imaginary ensemble representing our ignorance as to which of the two possibilities actually obtained. Unfortunately, the temporal evolution,

$$P_{\alpha\otimes\phi}\rightarrow|C_1(t)|^2P_{\alpha\otimes\phi}+|C_2(t)|^2P_{\delta\otimes\theta}$$

is absolutely impossible within the dynamical scheme of quantum theory, unless the composite system interacts with another system. However, it is pointless to multiply the number of systems, for at each state the same objections would arise, together with the same demand that an impossible total density operator is the "correct" one. Nor would it help to assume, as in Heisenberg's theory of measurement, that the cat-atom system, to be observed at all, must be immersed in an environment (Heisenberg's "external world") described by a mixture, say $\rho = \sum_R w_R P_{\gamma_R}$. It is true that the new total system--cat, atom, and surroundings--would then be in a mixed state at time t; but note closely its form (immediately derivable from the linearity of quantum dynamics):

$$P_{\alpha\otimes\phi} \otimes \sum_R w_R P_{\gamma_R} = \sum_R w_R P_{\alpha\otimes\phi\otimes\gamma_R}$$

$$\longrightarrow \sum_R w_R P_{[c_1^{(R)}\alpha\otimes\phi\otimes(\sum_n d_n^{(R)}\gamma_n) + c_2^{(R)}\delta\otimes\theta\otimes(\sum_m q_m^{(R)}\gamma_m)]}$$

Every component of the resultant mixture has a "blurred" cat in it!

Schrödinger's cat is of course a metaphor; what it represents is the notion of classical system, about which there are naturally many preconceptions. Chief among these is the cherished belief that a classical system connot take part in statistical considerations which include the so-called "interference" of probabilities which occurs for quantum states. A classical system always possesses a definite value for every classical observable, although there may be ignorance as to which value; but if so, the associated probabilities do not "interfere". One might ask: why so much interest in prequantum ideas? After all, there is no such thing as a classical system, except in a special limiting case of quantum theory. Besides, as already suggested, the cat paradox is based on the unwarranted association of ψ with a single cat-atom system in an almost occult sense. Thus a superposition of two eigenstates for a classical system is regarded as a kind of unreal, smeared representation which does not recognize that

at all times such systems possess either one eigenvalue or the other. We have already granted in section 4 that in the quantum framework, observables are not possessed; but it is a gross distortion to say that a superposition of eigenstates represents anything "blurred". Consider again the two density operators for the cat-atom system:

(1) $\rho = P_\psi$, $\psi = c_1 \alpha \otimes \phi + c_2 \delta \otimes \theta$

(2) $\rho = |c_1|^2 P_{\alpha\otimes\phi} + |c_2|^2 P_{\delta\otimes\theta}$.

The physical meaning of (1) is just this: in an ensemble of cat-atom systems examined at time t, the fraction $|c_1(t)|^2$ of the systems will display a live cat (and unchanged atom); the fraction $|c_2(t)|^2$ will exhibit a dead one (and radioactive decay products). Moreover, (2) means exactly the same thing so far as the observables in question are concerned. As has already been discussed at length, it is improper to regard pure states as referring to single systems and mixed states to imaginary ensembles expressing ignorance. Every density operator, pure or mixed, has the same referent--an ensemble.

This is not to say that (1) and (2) are identical; in principle, there exist observables whose measurement statistics are different in the two cases; hence the only scientific way to show that (2) is preferable to (1) would be to study empirical measurement results for such an observable. In the absence of such evidence, there is no reason to prefer (2) to (1) provided quantum theory is understood, not in the Copenhagen interpretation, but rather as outlined in sections 2-5. To insist that a composite system which is partly "classical" cannot be in a superposition of eigenstates is therefore quite dogmatic.

Nevertheless, many theories of measurement differ from the standard one given in section 6 by imposing the additional requirement upon

that it be classical.* As a result, much effort is expended in formulating reasons for replacing the inevitable post-measurement pure state of $S+M$ by a mixture. For Pauli measurements of the first kind, this means that the expression, $T_A(\psi \otimes \chi_0) = \sum_k \langle \alpha_k, \psi \rangle \alpha_k \otimes \gamma_k$, has to be reconciled somehow with the "real, classical" post-measurement density operator, which is taken to be the incoherent superposition, $\hat{\rho} = \sum_k |\langle \alpha_k, \psi \rangle|^2 P_{\alpha_k \otimes \gamma_k}$. (The S -ensemble itself still undergoes the von Neumann measurement transformation, $P_\psi \longrightarrow Tr_2 \hat{\rho}$ $= \sum_\ell \langle \gamma_\ell | \sum_k |\langle \alpha_k, \psi \rangle|^2 P_{\alpha_k \otimes \gamma_k} | \gamma_\ell \rangle = \sum_k |\langle \alpha_k, \psi \rangle|^2 P_{\alpha_k}$, just as in the standard theory.) Similar conversions to mixtures supposedly "should" accompany any other correlation scheme usable for measurement, owing to this supposed "classical" nature of M. Thus the Landau and Lifshitz correlation, $T_A(\psi \otimes \chi_0) = \sum_k \langle \alpha_k, \psi \rangle \beta_k \otimes \gamma_k$, would be supplanted by $\hat{\rho} = \sum_k |\langle \alpha_k, \psi \rangle|^2 P_{\beta_k \otimes \gamma_k}$.

The means employed to achieve this goal are often interesting, even though the end itself is of dubious value. For example, Landau and Lifshitz[50] define "classical" as follows: "The classical nature of the apparatus appears in the fact that at any given instant, we can say with certainty that it is in one of the known states" γ_n. Then, applying this criterion to the expression,

$$T_A(\psi \otimes \chi_0) = \sum_k \langle \alpha_k, \psi \rangle \beta_k \otimes \gamma_k,$$

they deduce that "the state of the system $[S+M]$ after the measurement will in actual fact be described, not by the entire sum, but by only the term which corresponds to the [apparatus 'reading' C_n]". It is remarkable that earlier in the same book, these authors assert that "it must be most

---

*Obviously, in this version of measurement theory, M includes the macroscopic laboratory equipment which the physicist ultimately "measures" by direct perception; thus in the Franck-Hertz experiment, M cannot be just the probing electron but must incorporate the complex devices with which the experimenter detects it and measures its energy.

decidely emphasized that we are here not discussing a process of measurement in which the physicist-observer takes part";[51] thus, like all Russian physicists, they reject at the outset the subjective tendencies of the Copenhagen orthodoxy. Yet their theory of measurement involves wave packet reduction! It is true that the reduction appears to follow from an objective principle, the requirement that $T_A(\psi \otimes \chi_0)$ be an eigenvector of C; but this is mathematically absurd. $T_A(\psi \otimes \chi_0)$ is a unique vector. It cannot be "in actual fact" different eigenvectors of C in different occurrences of the measurement process it represents. Apparently what Landau and Lifshitz have done is to postulate implicitly a second kind of state change, an objective wave packet reduction* which occurs under conditions which would be identified as a measurement situation by an alert physicist who happened to be present. Nevertheless, their theory, like that of Blochintsev (section 8), is in effect equivalent to the Copenhagen version in its physical meaning and is therefore subject to the critique given in section 8.**

Another method of justifying replacement of the pure state by a mixture is to define some sense in which the two are equivalent; the definition would also serve to identify the "classical" level within a quantal context. But in what sense can two unequal density operators, $\rho^{(1)}$ and $\rho^{(2)}$, be physically "equivalent"? Clearly they are distinguishable only by comparison of the measurement statistics they entail. Thus $\rho^{(1)}$ and $\rho^{(2)}$ are

---

*It might seem that if M were in a mixed state, a sort of "deterministic wave packet reduction" could be derived, with different "reductions" corresponding to interactions with the various components of the M -mixture. Thus "ignorance" of M would be blamed for the probabilistic character of quantum predictions. That this is impossible, at least for Pauli measurements of the first kind, has been shown by von Neumann,[52] Wigner,[53] and Komar.[54]

**The "projection postulate" of Landau and Lifshitz is somewhat generalized, but the principle is unchanged.

are certainly equivalent if, for every A, $Tr(\rho^{(1)}A) = Tr(\rho^{(2)}A)$; in fact this is true in the extreme case when $\rho^{(1)} = \rho^{(2)}$. If, on the other hand, only a restricted set of operators, $\mathcal{A} \equiv \{A_R\}$, is considered, $\rho^{(1)}$ and $\rho^{(2)}$ will be indistinguishable relative to $\mathcal{A}$-measurements, provided $Tr(\rho^{(1)}A_R) = Tr(\rho^{(2)}A_R)$, for every A_R in the set $\mathcal{A}$. This concept, which we shall call $\mathcal{A}$-equivalence and denote by $\rho^{(1)} \underset{\mathcal{A}}{\sim} \rho^{(2)}$ is occasionally used to secure the desired post-measurement mixture.

For definiteness, consider Pauli measurements of the first kind; the problem of measurement, for theorists worried about the "classical" aspect of $\underline{M}$, is now reduced to the following: find a meaningful restriction to place on $\mathcal{A}$ such that $\rho^{(1)} \underset{\mathcal{A}}{\sim} \rho^{(2)}$, where $\rho^{(1)} = P_{\sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \tau_R}$, $\rho^{(2)} = \sum_R |\langle \alpha_R, \psi \rangle|^2 P_{\alpha_R \otimes \tau_R}$. The observables corresponding to operators in $\mathcal{A}$ are then called the "classical" ones, i.e., those directly apprehended by the laboratory physicist, who cannot therefore distinguish $\rho^{(1)}$ from $\rho^{(2)}$. Examples of quantum measurement theories in which $\mathcal{A}$-equivalence plays this role are those of Feyerabend,[55] Wakita,[56] and Jauch;[57] but these are motivated more or less by an understanding of basic quantum theory in which Schrödinger's cat allegory is a <u>paradox</u>. We have already considered this position and dismissed it as an unfortunate byproduct of the Copenhagen interpretation.

However, such investigations do have merit as checks on the universal applicability of quantum theory. In this sense, they have little to do with measurement theory but fall rather into the realm of the correspondence, or "classical-limit", problem. It is reassuring to know, for example, that the quantal analysis of an interaction between a cosmic ray and a billiard ball will not yield predictions about the latter which are observably false by classical experiments. Whether the interaction has anything to do with measurement is irrelevant. Laboratory apparatus accounts for a

relatively small segment of the vast domain of "classical" objects.

In the measurement case, as already mentioned, both $\rho^{(1)}$ and $\rho^{(2)}$ imply that M is in the same mixed state: $\rho_2 = Tr_1 \rho^{(1)} = Tr_1 \rho^{(2)} = \sum_R |\langle q_R, \psi \rangle|^2 P_{\gamma_R}$. Hence, relative to M-observables, there is no difference between $\rho^{(1)}$ and $\rho^{(2)}$. If the operator C is taken to represent "meter readings", a classical observable, then $\rho^{(1)}$ certainly does not suggest that these readings "interfere" with each other in any way which could contradict classical observations of pointers. In fact, if the only observable classically measured on M were its "pointer position", then its measurement statistics could just as well be represented by the superposition $\psi = \sum_R \langle q_R, \psi \rangle \gamma_R$, since $\rho_2 \simeq P_\psi$; even this would not contradict classical experience! We shall return in the next section to this question of the significance of the concept "classical" in quantum physics.

The main point of the present section has been just this: the fact that M is in some sense "classical" does not imply that in all measurement processes a density operator like $\rho^{(2)}$ must be the "actual" result. That assumption far transcends the authority of classical physics in its role as a limiting case of quantum theory, and it even leads ultimately to strictures on the observability of those operators falling outside the preferred $\mathcal{A}$-equivalence class. Such arbitrary restrictions on the form of ρ are merely intuited along the lines of the Schrödinger cat "paradox", and do not relate to genuine "classical" experience; they are drawn instead from unjustified extensions of classical rules, extrapolations apparently motivated by the Copenhagen interpretation.

14. On the Proper Role of the Concept "Classical" in Quantum Physics

In discourses on complementarity, Bohr repeatedly insisted that classical description plays a role in quantum theory which is unavoidable and

of fundamental significance. His often quoted declaration, "...however far the phenomena transcend the scope of classical physical explanation, the account of all evidence must be expressed in classical terms",[58] has been echoed again and again. Heisenberg, for example, notes that the language of the laboratory employs the concepts of classical physics, and then asserts that "we cannot and should not replace these concepts by any others".[59] Such remarks go far beyond the milder and more reasonable asymptotic requirement that no quantal prediction concerning classically describable "macroscopic experience" should contradict valid classical prediction. Strictly, of course, there is always theoretical contradiction in the sense that quantal and classical constructs are quite different*; the correspondence principle can require only empirical agreement. This, however, is not the point stressed in the foregoing quotations.

The concept "classical" is there given a more basic status in quantum theory; Landau and Lifshitz characterize its double role as follows: " quantum mechanics contains classical mechanics as a limiting case, yet at the same time it requires this limiting case for its own formulation."[63] Undoubtedly, it is the latter belief which underlies attempts to construct measurement theories of the type discussed in the preceding section, wherein consideration of the "classical" nature of apparatus is regarded as essential.

Are we logically forced to accept the claim that classical physics is the cornerstone of quantum theory? Must the language of the laboratory be classical? To answer the first question, contrast the correspondence postulate (P1) as presented in section 2 with the following popular formulation which does make classical mechanics appear to be the basis of quantum mechanics:

P1a: The observables q (position) and p (momentum) correspond to

*This point has been discussed at length by Bohm,[60] Feyerabend,[61] and Hanson.[62]

operators Q, P which satisfy $[Q, P] = i\hbar 1$. Any observable a, whose state function in classical mechanics is $a(q, p)$ corresponds to a Hermitean operator of the form $A = a(Q, P)$.

Note that the very concept of observable is here construed to be basically classical; quantal representatives of observables are generated from their classical analogues. However, the effectiveness of this procedure (which, incidentally, is not always logically consistent) is obviously limited by the fact that quantum theory considers observables for which no classical analogue is imaginable. Nevertheless, quantum field theory, for example, is often introduced "heuristically" or "inductively" by generalizing Pla to an unconvincing method called "quantization". The concept of quantum field is then induced from a bizarre analysis of classical continuum mechanics in which field strengths become, upon "quantization", noncommuting field operators.

Actually, Pla and its generalizations are not required at all among the basic principles of quantum theory. The notion that classical physics is the foundation of quantum physics has an evident historical origin, but is of no logical value. Both theories have the same epistemological status as verified connections among their constructs, which are related in well defined ways to the given, the data of empirical experience. However, for historical reasons and because quantal and classical accounts must be empirically compatible within the classical sphere of interest, many quantal rules of correspondence appear to be based on classical physics. Bergmann[64] made much of this in his "logic of quanta". Nevertheless, this is essentially a backward-looking position; the classical world view, properly understood, is not self-evident, nor is it forced upon us by percepts. Like quantum theory, its logical genesis was an act of scientific creativity, or construction. Thus it seems preferable to formulate the correspondence

postulate (P1) as in section 2, a statement which recognizes no logical dependence of quantum theory upon classical physics. The "quantization" process (P1a) is then diminished to its correct status as a mnemonic device sometimes useful to classically trained physicists.

As for the second question, from the same philosophic perspective, it is clear that classical physics need not and perhaps ultimately should not be the standard mode of experiment description. The reasonable assertion that laboratory procedures be reported in communicable, "common-sense" language simply does not imply what Bohr and Heisenberg suggest, viz., that whatever experimental operations are performed must be described classically. Consider, for example, the complex of sensations which we categorize as the "motion of a Maxwell top" (an antique device seemingly as "classical" as anything could ever be). The primitive datal percepts involved are certainly neither classical nor quantal; moreover, these terms are not necessarily applicable to the empirical constructs used to describe and quantify observations and results of operations on the top. Only the far more abstract constructs and their interconnections which comprise the physical theory created to explain these empirical observations can be reasonably called classical or quantal. However, when a given theory is well entrenched, this "epistemological depth" of its constructs is forgotten in practice, and experiments come to be described in abstract terms provided by the theory itself. In the case of Maxwell's top, an empirical fact of interest is the variation in wobbling patterns which accompanies adjustments of the screws on the sides of the top; but a complete report of this observation is communicable without the sophisticated concepts of any physical theory, although such a description would be cumbersome and verbose indeed. But since classical mechanics provides the established theory of the top, the changing patterns of its motion occasioned by screw

adjustments may well be described in terms of "observations of the dependence of angular velocity upon the inertia tensor"--a truly "classical" laboratory language. Hence familiarity with a successful theory (classical mechanics) has created the illusion that its profound constructs are directly perceptible or self-evident; i.e., the "classical" laboratory language comes to be regarded as necessary, an unfortunate epistemological mistake. After all, quantum theory, too, can fully explain the observed wobblings of the top; and it could even provide a "quantal" laboratory language, familiarity with which can, and perhaps some day will, lead to its adoption as the "necessary" vernacular of common sense description. We therefore reject the principle that the perceived world is somehow inherently "classical" and that the quantum theory of measurement must have a "classical" aspect.

One of the trends in this kind of measurement theory was examined in the preceding section. The approach reviewed there was rather formal, the method being to "define away" allegedly undesirable interference terms. Another way to secure the desired "classical" aspect is associated with the names Jordan[65] and Ludwig,[66] who advocate thermodynamic analysis of the measuring apparatus. (In a similar vein, Wigner[67] has suggested that the WAY theorem proves $\underline{M}$ cannot be microscopic, hence its macroscopic character must be considered in the theory of measurement.) However, again the underlying purpose is apparently to derive the von Neumann measurement transformation, which is, as we have seen before, the recurrent goal of most measurement theories. An elaborate attempt along these lines due to Daneri, Prosperi, and Loinger[68] has been endorsed by Rosenfeld,[69] an outspoken apologist for Copenhagen ideas (Bohr's in particular). In their theory, the measurement transformation is derived by expressing the "classical" nature of apparatus in terms of ergodicity conditions and defining

macro-observables to be temporal averages of quantal observables. It is then shown that the quantal dynamics of $S+M$, if supplemented by these conditions, effectively yields the measurement transformation and explains the registration of a permanent "reading" in M. According to Rosenfeld,[69] "The main purpose of the analysis of measurement is to exhibit the physical process to which this formal 'reduction' [the measurement transformation] corresponds", and this Daneri-Prosperi-Loinger theory fulfills that requirement.

Probably such a demonstration does offer an approximate explanation of some actual measurement schemes; but as we have already observed, its basic structure is the derivation of an unnecessary, even rare, property (the measurement transformation) from an erroneous metaphysical belief (that apparatus is inherently "classical"). Clearly this approach can no more claim to be the quantum theory of measurement than could the variations of the standard theory considered earlier.

15. Infinite Regression

Thoughtful analysis of the standard theory of quantum measurement, or any of its variations, leads most theorists to recognize an interesting basic property of the usual approach. This property, sometimes called infinite regression, is received with varying degrees of enthusiasm depending on the metaphysical outlook of the critic. The essence of infinite regression is contained in this question: what performs the measurement upon M? Ordinary measurement theory can only reply that a second apparatus M_1 must interact with M in the same manner M interacts with S, i.e., with the effect that a measurement performed upon M_1 permits certain prediction of what a concurrent measurement on M would have yielded. Obviously, this suggests inquiry as to what makes measurements on M_1, and

so on ad infinitum.

In his original formulation of standard quantum measurement theory, von Neumann did not regard infinite regression as an undesirable attribute, but rather as a necessary characteristic expressing in mathematical terms the notion of psycho-physical parallelism. This idea derives from the elementary principle that all empirical observations must ultimately be regarded as perceived by the mind; the perception itself is an utterly primitive awareness of the given, a process intrinsically irreducible to scientific law. Thus, in every application of the scientific method, at some stage there must be statements to the effect that an observer simply observed some datum, and this will be true no matter how far into his brain the scientific analysis penetrates. Consider, for example, a measurement apparatus which registers its result as a pointer reading. It is most practical to terminate the analysis of this measurement act by saying that the observer observes the position of the needle. However, it is possible to go much further; for example, suppose the observation is made visually. An electrodynamic treatment of the relevant interactions among pointer, light, and eye can be involed to explain the formation of a retinal image of the needle and scale; but if this work is carried out in hopes of explaining away the observer, the effort is wasted. Instead of saying the observer observed the needle, we can now say that he observed the retinal image of it, but the necessity of the observing consciousness itself is as strong as before. It should be clear that no study of the optic nerve or even of electrical properties of the brain could possibly terminate otherwise than in a statement that the observer becomes aware of the needle position, or perhaps that this awareness occurs simultaneously with some electrical effect in his brain, which would mean that a neurophysiologist studying the observer's brain would observe, say, a certain

electroencephalogram pattern concurrently with the observer's announcement of the needle observation.

The primacy of the conscious mind in all scientific endeavor has the character of a general philosophic truth, and it is unfortunate that this lofty point was ever dragged down even as close to practical physics as quantum measurement theory. There the false impression has arisen that physics, or at least quantum physics, possesses an undesirable subjective element which must be reckoned with somehow.

In attempting to mathematize the subjectivistic excesses of some Copenhagen pronouncements, von Neumann therefore proposed two distinct processes, motion (P3) and measurement (P4), the latter representing the final transition to a consciousness. His motivation for drawing up the standard theory of measurement was to establish the consistency of P3 and P4 in the sense that the "cut" between observer and observed which P4 bridges should be arbitrary. Thus, in the standard theory, the same results obtain for S if M measures S or if M_1 measures $S+M$, etc.

However, we have seen earlier that this theory cannot reasonably be called the quantum theory of measurement. Are we therefore faced with an unusual subjective feature in quantum theory? The answer is negative, for in light of the understanding of quantum theory elaborated in foregoing sections, we deny not just the popular solution to this "quantal mind-body problem" but the problem itself. The foundations of quantum theory nowhere exhibit any more or less "subjectivism" than does classical mechanics; both theories, as has already been noted in another context, are easily accomodated by the same epistemological framework. And infinite regression is as much a property of classical as of quantum theories.

Nevertheless, von Neumann's recognition of that property and his mathematical enshrinement of it in the projection postulate has sometimes

induced the belief that quantum theory carries a destructive subjectivistic quality which must be eliminated in order to save objective science. Jauch, for example, offers this as a reason that justification must be found for replacement of pure states involving apparatus by mixtures (section 13); presumably, this would halt the regression by inserting a classical level and therefore closing out the unwanted subjectivism. This illusion has its roots in the mechanistic philosophy widely held in the classical epoch of physics, when physical laws were widely regarded as purely objective "discoveries" totally divested of any metaphysical format constructed by the physicists themselves. Actually, this tenet was philosophically unacceptable[70] even in the heyday of classical physics; it is therefore strange that objectivity for quantum theory should be sought by relating it to classical physics. Both theories are subjective and objective in exactly the same ways,[71] and both display the same infinite regression property. The main point we wish to emphasize here is that this characteristic is not problematical, does not deprive science of objectivity, but rather indicates that objectivity is established within "subjective" experience.* However, the problem of measurement in quantum physics in the context of the present investigation is not of such philosophic depth as to require further discussion in this vein.

Accordingly, we now dismiss this basic notion of infinite regression from further consideration, since it darkens more than it illumines the problem at hand, viz., to clarify the meaning of the quantal terms measurement and preparation. Nevertheless, the logical structure of the infinite regression analysis does prove to be quite valuable in this connection, provided the above mentioned efforts to link it to the mind-body problem

*For a further discussion on objectivity in quantum mechanics, cf. H. Margenau and J. L. Park, Delaware Seminar on the Philosophy of Science III, Mario Bunge, Ed., Springer-Verlag (1967), Ch. 8.

are forgotten.

Consider again the skeletal framework of quantum measurement theory, according to which an a-measurement on S consists of an interaction with an a-meter $M(a)$ which establishes some correlation between relevant states of S and M. The a-measurement is then carried out by observing the "reading" of $M(a)$. Undoubtedly, this account does offer the correct quantal description of many laboratory procedures; but we now suggest that it does not deserve the name usually given it--the quantum theory of measurement. This "theory" cannot be said to explain the concept measurement; indeed, as we shall see below, this theory cannot even be stated carefully without implicitly using the term measurement itself several times. In this respect, a quantal description of a measurement process differs markedly from its classical counterpart, which does not require the term measurement at all until the final stage when an observer "looks at" the meter. We shall see below that the resultant dichotomy of meaning for the term measurement in its classical and quantum usages is traceable to the respective characters of classical and quantal observables (section 4).

To verify our claim that so-called quantum measurement theory is not even statable without using the term measurement itself, consider its essential feature, the establishment of correlations. In classical physics, where observables may be assigned values possessively, correlations between S and M refer to these possessed physical quantities independently of measurement. This scheme, however, is inconceivable within the quantal framework, owing to the essential latency of observables. No matter what specific form correlations may take, in quantum theory they are inevitably nothing but connections among potential measurement results. For example, the correlation assumption, $T_A(\psi \otimes \chi_0) = \sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \gamma_R$,

strictly implies only this: a simultaneous a -measurement$_1$* on S and c -measurement$_1$ on M at the completion of the measurement$_2$ interaction will yield the pair (a_k, c_ℓ) with nonzero probability only if $k = \ell$. It is therefore said that the measurement process renders the a -measurement on S redundant, since a c -measurement on M is sufficient for prediction with certainty as to what the post-measurement$_2$ a -measurement$_1$ would yield. (It is interesting to note further that nothing can be said with certainty about what result would be obtained in an a -measurement$_1$ on S just before the measurement$_2$ interaction, except that $|\langle a_k, \psi \rangle|^2$ is the common probability distribution for a -measurements$_1$ before and after measurement$_2$ interactions of this type.)

Thus we see that a rigorous quantum description of a measurement correlation process is a verbally cumbersome account in which the concept measurement itself enters repeatedly. This recurrent use of the term measurement is unavoidable in any quantal description which adheres strictly to the latent character of quantum observables. It will be urged in the concluding sections that this essential recurrence is the key to understanding the epistemological status of the quantum term measurement. To summarize: the goals of the present section have been (1) to point out that the philosophic problem of infinite regression to consciousness is equally relevant to both classical and quantum physics (and equally beyond the proper domain and competence of both); and (2) to show that close logical scrutiny of any measurement scheme in a manner suggested by the infinite regression argument (viz., posing questions like "in what sense does M measure S ?" and "what measures M ?") reveals that the quantum concept measurement, unlike the classical one, must appear as a primitive term even in the so-called quantum theory of measurement itself.

*Subscripts on the term measurement will be referred to later.

16. Quantum Explanation of a Real Measurement

The favorite motivating experiment for measurement theorists seems to be that of Stern and Gerlach, which is alleged to be an example of the correlation $T_A(\alpha_R \otimes \chi_0) = \alpha_R \otimes \gamma_R$, and is occasionally elevated to the status of prototype for most, if not all, measurements. This view clearly exaggerates its importance; nevertheless, the Stern-Gerlach experiment is a good one to examine, if only because of its relative simplicity. We therefore present a somewhat unconventional analysis of it.

Before getting immersed in the mathematics, let us briefly recapitulate the data originally reported by Stern and Gerlach. A beam of silver atoms, emanating from a slit in a furnace, was channeled between magnetic pole pieces toward a glass plate, upon which silver deposits eventually accumulated. One pole piece was knife-edged, the other flat; hence, the silver atoms traversed an inhomogeneous magnetic field. Stern and Gerlach studied microphotographs of the deposits and interpreted what they saw as follows: "The pictures show that the silver atom beam in an inhomogeneous magnetic field is split up into two beams in the direction of the inhomogeneity, one of which is attracted to the knife-edged pole and the other of which is repelled."[72] This 1922 description is slightly tainted by classical language. A "pure" quantum theorist would interpret the same photographs this way: Position measurements on an ensemble of silver atoms, each prepared by emission from a furnace and passage through an inhomogeneous magnetic field, yield results whose statistical distribution exhibits two sharp peaks along the direction of inhomogeneity of the field. (Often there are more such peaks, but if ground state hydrogen atoms are used, as Phipps and Taylor[73] have done, there are always just two.)

To explain this phenomenon quantum mechanically, the initial state vector of the hydrogen atom* upon emergence from its source is assumed to

*We take for simplicity the Phipps-Taylor case.

be of the form $\psi \otimes \chi_0$. In the Schrödinger-Pauli representation, the spinor ψ involves only electronic coordinates relative to the nucleus, while χ_0 is a fairly localized wave packet whose argument is the atomic "center of mass". Thus the atom is formally regarded as though it were a composite system whose constituents are initially in states ψ and χ_0, a feature to be exploited later on (section 19). Let α_1, α_2 be the eigenvectors belonging to the component of spin in the inhomogeneity direction of the magnetic field. When the temporal evolution from initial state $\psi \otimes \chi_0$, where ψ is the ground state, is calculated, the following result is obtained:

$$T(t)(\psi \otimes \chi_0) = \sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \gamma_R(t) \equiv \Psi(t).$$

Of special interest is the center-of-mass motion, represented in the equation above by $\gamma_R(t)$, since it is the final position distribution of the atoms that the Stern-Gerlach apparatus displays. This problem is solved by examining the final center-of-mass position probability density. If δ_{xyz} denotes a common eigenvector of center-of-mass coordinates x, y, z, the required probability density is

$$\begin{aligned} w(x,y,z;\Psi(t)) &= \langle \Psi(t), 1 \otimes P_{\delta_{xyz}} \Psi(t) \rangle \\ &= \Big\langle \sum_\ell \langle \alpha_\ell, \psi \rangle \alpha_\ell \otimes \gamma_\ell, \sum_R \langle \alpha_R, \psi \rangle \alpha_R \otimes \langle \delta_{xyz}, \gamma_R \rangle \delta_{xyz} \Big\rangle \\ &= \sum_{R\ell} \langle \psi, \alpha_\ell \rangle \langle \alpha_R, \psi \rangle \delta_{\ell R} \langle \gamma_\ell, \delta_{xyz} \rangle \langle \delta_{xyz}, \gamma_R \rangle \\ &= \sum_R |\langle \alpha_R, \psi \rangle|^2 |\langle \delta_{xyz}, \gamma_R \rangle|^2 \end{aligned}$$

It now turns out[24] that if z is the direction of field inhomogeneity, $|\langle \delta_{xyz}, \gamma_1 \rangle|^2$ is negligibly small except in the same z-interval as one of the observed accumulations on the final plate; similarly, $|\langle \delta_{xyz}, \gamma_2 \rangle|^2$ practically vanishes outside the neighborhood of the second deposit. The theory therefore fully accounts for observations of the Stern-Gerlach type described empirically above. Moreover, the theory also reveals an interesting correlation between the internal eigenstates and the center-of-mass

motion. To be specific, in the expression for $w(X, Y, Z)$, the "strength" of the kth peak $|\langle \delta_{xyz}, \gamma_k \rangle|^2$ is "weighted" by $|\langle \alpha_k, \psi \rangle|^2$, a functional of the internal eigenvector α_k. This property is often invoked in Stern-Gerlach-centered discussions on measurement theory. However, before getting into that, let us determine to what extent the quantal concept measurement has already been used in the foregoing theoretical explanation of the actual Stern-Gerlach data.

First an assumption was made about the initial state $\psi \otimes \chi_0$; this amounted to a number of conditional statements involving measurements never performed. For example, to assume the hydrogen atom is initially in its ground state means, among other things, that if energy were measured, the result would be the lowest energy level. This only illustrates that the concept of preparation ultimately depends on that of measurement. To say that a certain physical act prepares a state ρ always implicitly entails a set of conditional statements involving measurement in an essential way. Nevertheless, in the Stern-Gerlach experiment itself, none of these measurements relating to the preparation of $\psi \otimes \chi_0$ is performed. It may be assumed that such measurements have been made extensively in the past on a similar oven-slit device or other source and that it is guaranteed to be a bona fide producer of ensembles with $\rho = P_{\psi \otimes \chi_0}$.

The only measurements mentioned above in connection with the Stern-Gerlach experiment as if they were actually performed are of the observables X, Y, Z, i.e., atomic (center-of-mass) position coordinates. By contrasting this description involving atomic position measurements to the more prosaic laboratory report of Stern and Gerlach, we obtain a first indication of a point to be developed later, viz., that the quantal concept of measurement is far more abstract and less empirical than its name suggests.

The original Stern-Gerlach detection scheme--microphotographs of silver deposits--was in fact too crude to perform a single position measurement. Yet quantum theory explains the pattern of silver deposits as if they represented numerous elementary position measurements, although the adherence of a single silver atom to the glass plate is certainly never really observed. However, position measurements upon single microcosmic systems are not impossible; on the contrary, position is in a sense the most nearly "observable" micro-observable there is, as will become increasingly evident below. Now, suppose the glass plate is replaced by a better detector which is able to perform an operation worthy of the name position measurement. For example, impact of a single atom may trigger an "avalanche" of reactions about the collision point which produce a photographable "spot". Atomic position can then be defined operationally by equating the center coordinates of the spot with the "result of a position measurement". (These coordinates are determined by a "ruler", a macroscopic device which, used correctly, will yield the same numbers regardless of the intuitive world view of the experimenter; indeed he may employ quantum, Newtonian, or Aristotelian mechanical concepts for his own personal thoughts about rulers.) Using that rule of correspondence to relate the construct position measurement to empirical observation, the quantal explanation of the Stern-Gerlach effect in terms of "single position measurements" is no longer problematical; however, this has not really been the main point of this paragraph. Of more general value is the identification of one rule of correspondence between a quantum observable (position) and a laboratory operation.

Despite its innocent appearance, the foregoing operational definition is in experimental practice not merely a specialized example; it is rather the fundamental rule of correspondence in quantum physics, in the sense

that all other quantum observables are actually measured by establishing correlations with the observable position. All measurement paraphernalia --photographic emulsions, cloud chambers, bubble chambers, counters-- "directly" measure position in a manner similar to that described above. As Lande[25] puts it, "...nowhere in physics do we have 'direct' data, the only exception being location in space and time, that is (q,t)-values. Velocity, momentum, energy, etc. are always determined indirectly." DeBroglie[76] makes the same point as follows: "Any process of measurement of a dynamic variable, such as the energy and momentum of a particle, is a complex and indirect process which necessarily utilizes direct observation of particle localizations." Probably deBroglie's "necessarily" is too strong; the dominant practical role of the observable position is a matter of fact rather than logic; but recognition of this fact sheds more light on the nature and meaning of the quantal construct measurement than do any of the so-called "theories of measurement" reviewed in previous sections. What it suggests is that most quantum observables are never "observed", and that most of the measurements which are unavoidably mentioned in every quantum theoretical explanation are in fact never performed. Indeed, in a certain sense, they are perhaps unperformable. The remainder of this work is devoted to the amplification and clarification of these remarks.

17. Construction of an Operational Definition

Consider the quantum observable spin. How can it be measured? What does it mean to say that an ensemble with state vector ψ will upon measurement of the Z-component of spin, S_z yield $\frac{1}{2}\hbar$ with relative frequency $|\langle \alpha_1, \psi \rangle|^2$? It is instructive to take a close look at how an operational definition of S_z is usually developed from the Stern-Gerlach experiment. Once again we suppose that an ensemble of ground state hydrogen atoms is

available for study. As noted above, the Stern-Gerlach apparatus brings about this state evolution: $T(t)(\psi \otimes \chi_0) = \sum_k \langle \alpha_k, \psi \rangle \alpha_k \otimes \gamma_k(t)$. We have already seen that w(X,Y,Z) suggests an interesting correlation between S_z and Z, due to properties of the γ_k. This may be seen more clearly by computing the joint probability for results of S_z and Z *-measurements at time t. Let $\{\Delta_k\}$ denote eigenvalues of S_z; then the joint probability density $w(\Delta_n, Z)$ is found as follows:

$$w(\Delta_n, Z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\langle \sum_\ell \langle \alpha_\ell, \psi \rangle \alpha_\ell \otimes \gamma_\ell \middle| P_{\alpha_n} \otimes P_{\delta_{xyz}} \middle| \sum_k \langle \alpha_k, \psi \rangle \alpha_k \otimes \gamma_k \right\rangle dX dY$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{k\ell} \langle \psi, \alpha_\ell \rangle \langle \alpha_k, \psi \rangle \langle \alpha_\ell, \langle \alpha_n, \alpha_k \rangle \alpha_n \rangle \langle \gamma_\ell, P_{\delta_{xyz}} \gamma_k \rangle dX dY$$

$$= |\langle \alpha_n, \psi \rangle|^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\langle \delta_{xyz}, \gamma_n \rangle|^2 dX dY.$$

Recall that $|\langle \delta_{xyz}, \gamma_n \rangle|^2$ almost vanishes except near one of the Stern-Gerlach accumulations, which we shall call the n th region. The distribution $w(\Delta_n, Z)$ therefore implies that with near certainty an S_z-measurement would yield Δ_n when and only when a simultaneous Z-measurement yields a result in the n th region. This leads to the common identification of the Stern-Gerlach apparatus as a kind of "spin-meter" which operates as follows: to measure the observable S_z (on a ground state hydrogen atom), direct the atom through a magnetic field inhomogeneous in the Z direction, and then measure Z, already operationally defined. A Z-result in the nth region is considered to be a "reading" Δ_n of the "spin-meter".

It is therefore tempting just to regard this procedure as the empirical meaning of the quantal term spin-measurement. Unfortunately, this cannot be done for two reasons: (1) the operational definitions of Z and S_z would then be contradictory, and (2) the Stern-Gerlach method cannot <u>be</u> a spin-measurement because its own detailed quantum mechanical description involves the concept spin-measurement in a logically anterior way. Both

*We now drop the notational distinction between observables S_z, Z and operators S_z, Z.

(1) and (2) follow from basic principles enunciated in previous sections.

Reason (1) is essentially a consequence of the WAY theorem, which in this instance prohibits γ_1, γ_2 from being orthogonal. Thus, although $|\langle \delta_{xyz}, \gamma_n \rangle|^2$ is minuscule outside the n th region, it does not vanish. Hence, there is a finite probability, for example, that simultaneous Z- and S_z-measurements would yield a Z-result in region 1 and the S_z-eigenvalue Δ_2. In other words, we are able to evaluate "how good" a "spin-meter" the Stern-Gerlach device is; therefore it cannot be used to define the quantal term spin-measurement. If it were so employed, the operational definition of Z would be contradicted: the appearance of a "spot" in the n th region would always mean that an S_z-measurement has yielded Δ_n but would no longer indicate with certainty that the Z-measurement result coincided with the "spot"! The best conclusion seems to be that the Stern-Gerlach "spin-meter" is excellent but not perfect, and hence unsuitable for defining the concept of spin-measurement. The importance of this result lies in the fact that in practice the above "spin-meter" seems to be the only kind there is; therefore, the construct spin-measurement--of proven value in theoretical explanations--refers to no actual "laboratory measurement" at all. This suggests perhaps that quantum physics uses the term measurement in two distinct senses, one traditional and one peculiarly quantal. That such is the case will emerge presently from the following consideration of reason (2).

Since it is independent of the WAY theorem and universally applicable to any conceivable quantal description of a "laboratory measurement", reason (2) is more fundamental than reason (1). The principal point has already been discussed in some generality at the end of section 14: the term measurement necessarily occurs as a primitive even in a quantal description of a measurement process. In the present case, this means

that a careful account of the operation of a Stern-Gerlach "spin-meter" runs as follows: if an S_z-measurement$_1$* on the atom just prior to the measurement$_2$ interaction with the "spin-meter" would certainly have yielded Δ_n, then immediately after the measurement$_2$ interaction, an S_z-measurement$_1$ will certainly yield Δ_n and a Z-measurement$_1$ will (almost) certainly yield a Z-value in the nth region. Hence the post-measurement$_2$ S_z-measurement$_1$ is redundant; a post-measurement$_2$ Z-measurement$_1$ is sufficient to deduce what an S_z-measurement$_1$ would have given at the instant the Stern-Gerlach measurement$_2$ procedure began. We hasten to point out that this is not mere semantic legerdemain. No alternative quantal description is conceivable; to explain in detail the operation of a Stern-Gerlach "spin-meter" in any other way is impossible within the language of quantum theory! Reference must be made to the imaginary results of S_z-measurements which are never performed in any laboratory. Furthermore, the very concept of performing an S_z-measurement seems to designate no empirical act whatsoever. Thus the Stern-Gerlach device is said to reveal "what an S_z-measurement would have given" earlier (just before the atom entered the magnetic field); yet this earlier S_z-measurement itself is not even an imaginable laboratory operation. Indeed, the very device which supposedly performs that S_z-measurement can itself be described only in terms of what the S_z-result would have been if S_z had been measured! As suggested in section 14, similar conclusions may be drawn from the quantal explanation of any "laboratory measurement" procedure whatever. An experimental scheme designed to "make measurements of a" will in general be described in terms of unperformed and unperformable a-measurements.

---

*As before, the subscripts should be ignored for the moment.

18. The Dual Meaning of Measurement in Quantum Physics

Imagine for a moment that S_z were an observable in the classical sense; it would then be meaningful to say that the atom entering the Stern-Gerlach device has $S_z = s_n$. Assume further that theoretical analysis demonstrates, in analogy to the quantal case, that such an atom is always channeled to the nth region. It would then be permissible to conclude that the term S_z-measurement merely refers to this operation: pass the atom through a Stern-Gerlach device and observe the spatial region of its emergence. This act is a measurement in the classical sense because it leads to a determination of what S_z-value the atom possessed. Furthermore, nowhere in this or any classical description of a measurement procedure does the term measurement itself enter in a fundamental way.

By contrasting this fictitious classical description of a Stern-Gerlach "spin-meter" to its quantal counterpart, the source of difficulty in the quantum case becomes apparent. At several points in the classical account of a measurement process, the concept of possessed observable is employed; but at the analogous places in a quantal account, this notion cannot be used. The basic structure of quantum theory forbids it. In section 4, we noted that this old concept of possession had been superseded in quantum theory by the idea of latency. Thus the quantum axioms (section 2) connect observables to systems and states only in a dispositional sense. This connection is made through the primitive construct A-measurement, about which nothing is said except that when it is performed upon a system, it yields a number. Hence the logically primitive construct A-measurement, a consequence of latency, plays a role in quantum theory analogous to that of possession in classical theory. Accordingly, to convert a classical description of a measurement procedure to a quantum description, each classical statement of the form, "S has $A = a_k$" must be replaced

by the quantal proposition, "an a-measurement upon S would yield a_k".

The term measurement, as it appears in the quantum axioms, has therefore a theoretical status quite distinct from that of the term measurement in its classical usage and in the phrase, "theory of measurement". In recognition of this homonymy, we shall henceforth designate the primitive construct measurement, which is essential to the statement of quantum axioms, as M_1; an a-measurement will be denoted by $M_1(a)$. On the other hand, $M_2(a)$ will represent the classical concept of measurement, or at least the nearest quantal analogue to it. For example, the fictitious classical $M_2(S_z)$ is an operation which employs physical interaction to establish a correlation between possessed S_z-values and Z-values of the atom, so that an observation of the Z-value (i.e., intelligently "looking at" the "spot" and "ruler") enables inference of the S_z-value. Similarly, the quantal $M_2(S_z)$ is an operation which employs physical interaction to establish a correlation between the potential results of $M_1(S_z)$ and $M_1(Z)$ so that the actual result of a performed $M_1(Z)$ enables inference of the potential result of a never performed $M_1(S_z)$. This important distinction between M_1 and M_2 was hinted at in sections 15 and 17, where subscripts were attached to the word measurement to suggest its two meanings in quantum parlance.

In connection with M_1 and M_2, several questions must be raised: (1) How does M_1 fit into the general epistemological framework of physics? (2) Similarly, what is the role of M_2 in the scientific method? (3) How are M_1 and M_2 related? and (4) What has all this to do with the quantum theory of measurement? We shall now discuss these points in that order.

(1) The most striking property of M_1 is its abstractness; it is an ultimate primitive construct irreducible to any others. Epistemologically it is like the concepts physical quantity and mass point in classical

mechanics--no phenomenon can be theoretically comprehended without it. Yet in spite of its deeply theoretical status, the nature of $\mathcal{M}_1$, as implied by its role in the quantum axioms, belies an abstract mimicry of a naive view which equates measurement and elementary observation. Thus the statement--"if $\mathcal{M}_1(a)$ is performed upon $\underline{S}$, it will yield the result a"--is beguilingly similar in form to this: "if an observation is made of the sky, it will 'yield' the color blue." However, the apparent similarity is purely grammatical; the differences are far more important. If a literal interpretation is demanded for the clause "$\mathcal{M}_1(a)$ is performed upon $\underline{S}$", then we shall have to provide some kind of "micro-elf", or "quantum demon", to do the performing; for, as we have just seen in the special case of spin-measurements, real physicists do not, indeed cannot, "perform $\mathcal{M}_1(S_z)$". $\mathcal{M}_1(S_z)$ is quite typical in this respect. Hardly any $\mathcal{M}_1(a)$ is ever "performed" by an experimenter; in practice, macro-position observations are perhaps the only exception. Nevertheless, the construct $\mathcal{M}_1(a)$, even if it is imagined to represent the perceptions of omnipresent q-demons, is invaluable and unavoidable in the quantum theoretical explanation of all actual empirical observations.

(2) In the opening remarks of section 1, measurement was introduced in the usual way as the epistemological link between percepts and concepts. On reflection we now see that $\mathcal{M}_2$ was the proper subject of that discussion. Moreover, in using $\mathcal{M}_2$ to introduce the essential use of the term measurement, in the quantum axioms, we confused it in the natural manner with $\mathcal{M}_1$, which we now recognize as the only measurement-construct appearing in the postulates. However, the fundamental mediatory role of $\mathcal{M}_2$ is the same in quantum physics as in the remainder of science; the novelty of the quantum framework lies in the fact that, as a consequence of the latency of observables, among the constructs which $\mathcal{M}_2$ relates to datal experience is $\mathcal{M}_1$.

(3) In classical physics, an $M_2(a)$ procedure is always understood as establishing a correlation between the possessed value of an abstract a and the possessed value of some observable χ directly accessible to the experimenter. With deBroglie and Landé, we have tentatively (cf. section 20) adopted the view that χ is always essentially a position; thus the physicist always "looks at" a "spot" and "ruler", thereby observing χ directly. It should be clear that without the latter direct observation, $M_2(a)$ would be impossible and the theory at hand therefore physically meaningless. Applied to the quantum case, this means that there must exist an χ directly observable by a real experimenter; i.e., the physicist himself must be the q-demon who performs $M_1(\chi)$ by a simple "look-and see" observation. As before, χ is presumably a position, which the physicist observes directly as a coincidence of "spot" and "ruler". Were he a quantum purist, he would of course describe his actions as follows: "$M_1(\chi)$ was performed ('ruler' placed near 'spot') and yielded x ('spot' coincided with 'ruler' mark x)." Thus for M_2 to be possible at all, a quantum physicist must for some a be himself a q-demon capable of "performing" $M_1(a)$, although for most a's his "performing $M_1(a)$" is as inconceivable as, for example, the direct perception of (possessed) energy by a classical physicist.

(4) Having established the twofold meaning of measurement in quantum physics, we are now able to state precisely the very most that any so-called quantum theory of measurement could hope to explain. Simply put, such a theory can only offer a description of an $M_2(a)$ in terms of $M_1(a)$ and other M_1's. To reach that conclusion, we assume that the final purpose of formulating a "quantum measurement theory" would be to give a quantal description of actual laboratory measurement processes, in particular, to achieve a quantum theoretical understanding of how information about the

microcosm is obtained. Roughly speaking, it is obvious that knowledge of things unperceivable must be gained through correlations with things directly apprehended. To be more scientific, a microsystem can be studied only via physical interaction with it; otherwise, the requisite correlations could not be established. On the other hand, a laboratory measurement scheme $\mathcal{M}_2$ can be exhaustively described without using any postulates except those normally required to explain other physical processes; this fact was essentially the crux of many previous sections in which we criticized the various extant ideas about measurement, some of which seemed to regard quantum measurement processes as more than just physical processes. Therefore, an $\mathcal{M}_2$ can and must always be explained in terms of the basic constructs of quantum theory, among which are the $\mathcal{M}_1$'s. The $\mathcal{M}_1$'s themselves, being ultimate primitive constructs, are not susceptive of further quantal explanation. A "quantum theory of $\mathcal{M}_1$" would be tautological, like a "mechanical theory of motion". Hence, a quantum theory of measurement can at most be a quantum theory of $\mathcal{M}_2$.

19. Remarks on Preparation

Several times in previous sections we have alluded to the interdependence of the concepts measurement and preparation. Unfortunately, many treatments of "measurement theory" fail to stress the differences between these concepts. Frequently, measurement and preparation are regarded as essentially equivalent; this premise leads inevitably to "measurement" discussions marked by severe ambiguities. For example, Schwinger's "algebra of measurement"[77] is really a hybrid "algebra of measurement and preparation" in which the two concepts are not carefully distinguished; and Groenewold[78] overtly ignores the difference: "I take all the time the term 'measurement' in a broad sense, including initial preparations, intermediate

observations and final detections. Those, who prefer another terminology, are free to make the translation."

This belief that measurement and preparation are prectically equivalent arises upon adoption at least of P4 as a universal quantum postulate. The latter has already been discussed at some length, the principal conclusion being that P4, although occasionally derivable, is required a priori in the quantal account of no phenomenon whatsoever. However, the equivalent treatment of measurement and preparation is most often founded upon the logically untenable predecessor of P4, which assigns a state vector to a single system on the basis of a measurement result. Thus it is sometimes asserted that an A-measurement which yields a_k prepares the state α_k.

In this context, it is difficult to say whether the A-measurement is $M_1(a)$ or $M_2(a)$. Very formal treatments sometimes give the impression that an $M_1(a)$ yielding a_k is equivalent to the preparation $\Pi(P_{\alpha_k})$ of the state α_k. Others seem to suggest that when $M_2(a)$ reveals the result a_k of an $M_1(a)$, $\Pi(P_{\alpha_k})$ may be regarded as having occurred at the time of $M_1(a)$, or even, as sometimes claimed,[79] at any time between $M_1(a)$ and the completion of $M_2(a)$! In our opinion such considerations are as nonsensical as the following sentence: "a q-demon prepares S in the state α_k by 'looking at' the A-ness of S and 'seeing' a_k."

A similar comment applies to the contention that a commitment regarding the post-measurement state is essential for the theoretical analysis of successive measurements. If "successive measurements" means successive M_1's, then the problem is unphysical; it amounts to an inquiry about a "sequence of q-demonic acts". On the other hand, if successive M_2's are contemplated, all that is involved is a physical process, fully describable without attributing any properties to successive M_1's.

Like $\mathcal{M}_2(a)$, $\Pi(P_{\alpha_R})$, or in general $\Pi(\rho)$, is correctly interpreted as a physical process which always has a quantum theoretical explanation. To illustrate this point and to clarify the distinction between measurement and preparation, we shall present briefly the careful quantal description of a preparation scheme, a quantum theory of preparation. Since it is commonly mentioned erroneously as evidence that measurement and preparation are the same, the Stern-Gerlach "spin-meter" as a preparation device will here be used to prove the opposite--that measurement and preparation are basically different.

Once again, for simplicity ground state hydrogen atoms will be fed into the magnetic field. This statement, as noted earlier, assumes that a preparation scheme for the initial ensemble of atoms is given. A theory of preparation can at most be a quantal description of the physical process by which a desired ensemble is transformed and/or extracted from an initial ensemble. It is of course impossible to describe the ensembles or the process without the primitive measurement construct $\mathcal{M}_1$. Only in this sense is preparation ultimately reducible to the concept measurement, but this is not the same as saying that measurement and preparation processes ($\mathcal{M}_2$ and Π) are equivalent.

If all members of the ensemble consisting of atoms which passed through a Stern-Gerlach field are considered, the device has obviously prepared the state, $T(t)(\psi \otimes \chi_0) = \sum_R \langle \psi, \alpha_R \rangle \alpha_R \otimes \gamma_R(t)$. However, since the position detector is absent, the preparation is not connected with any measurement operation--a simple illustration that Π and $\mathcal{M}_2$ are not equivalent.

More interesting preparations based on the Stern-Gerlach experiment require the concepts of subensemble selection and $\mathcal{A}$-equivalence. For definiteness, we consider $\Pi(P_{\alpha_1 \otimes \gamma_1})$, a preparation process often

associated with the Stern-Gerlach setup. The customary description of the method is rather naive: since the magnet has spatially separated the "beam" into the two disjoint regions, an $\alpha_1 \otimes \gamma_1$ -"filter" may supposedly be constructed by erecting an absorber W in region 2. The complete apparatus--$\Pi(P_{\gamma_2 \otimes \alpha_2})$-device, Stern-Gerlach magnet, and W--would therefore constitute a $\Pi(P_{\alpha_1 \otimes \gamma_1})$-device. Now, it may be that this experimental arrangement does indeed effect $\Pi(P_{\alpha_1 \otimes \gamma_1})$; but, if so, there must exist an explanation better than the preceding "filtration" argument. This notion of filtering arises from an erroneous classical interpretation of the aforementioned probability distribution $W(\alpha_1, z)$. As has been stressed repeatedly, quantum theory involves only probabilities that a specified measurement will yield a given result; it does not and cannot meaningfully speak of the probability that a system will be found "in" a given state. Thus the fact that $W(\alpha_1, z)$ practically vanishes outside region 1 does <u>not</u> mean that an atom detected in that region was "really in" the state $\alpha_1 \otimes \gamma_1$ all along, an obvious presupposition behind the above "filtration" scheme. The ensemble to be "filtered" is in fact pure --its density operator is the projection $P_{\sum_k \langle \alpha_k, \gamma \rangle \alpha_k \otimes \gamma_k}$ --and therefore <u>in principle</u> irreducible to distinct pure subensembles such as $P_{\alpha_1 \otimes \gamma_1}$.

This invites the possible reply in defense of the "filtration" picture that we have been ignoring the very practical fact that only a restricted set of observables $\mathcal{A}$ is really of any interest here. Thus our admonitions regarding the theoretic impossibility of subensemble selection, or filtration, (relative to <u>all</u> observables) may be irrelevant if just the observable set $\mathcal{A}$ is studied. In terms of $\mathcal{A}$-equivalence, it may be that

$$P_{\sum_k \langle \alpha_k, \gamma \rangle \alpha_k \otimes \gamma_k} \underset{\mathcal{A}}{\sim} \sum_k |\langle \alpha_k, \gamma \rangle|^2 P_{\alpha_k \otimes \gamma_k},$$

hence that the desired selection may actually be possible relative to $\mathcal{A}$.

As a matter of fact, if $\mathcal{A}$ is the set of all observables either of the form $A\otimes 1$ or $1\otimes B$, then the foregoing $\mathcal{A}$-equivalence relation is valid (to an excellent approximation)*. To prove that, recall that the atom is formally treated as a composite system in the sense that its "internal motion" is separated from its "center-of-mass motion"; thus the desired $\mathcal{A}$-equivalence may be established by proving the following relations:

(1) $Tr_1\rho = Tr_1\rho_M$,

(2) $Tr_2\rho = Tr_2\rho_M$,

where $\rho = P_{\sum_k \langle\alpha_k,\gamma\rangle \alpha_k\otimes\gamma_k}$, $\rho_M = \sum_k |\langle\alpha_k,\gamma\rangle|^2 P_{\alpha_k\otimes\gamma_k}$.

The required calculations are straightforward:

$$Tr_1\rho = \sum_n \langle\alpha_n|\sum_\ell \langle\alpha_\ell,\gamma\rangle\alpha_\ell\otimes\gamma_\ell\rangle\langle\sum_k\langle\alpha_k,\gamma\rangle\alpha_k\otimes\gamma_k|\alpha_n\rangle$$
$$= \sum_{n\ell k}\delta_{\ell n}\langle\alpha_\ell,\gamma\rangle\delta_{kn}\langle\gamma,\alpha_k\rangle|\gamma_\ell\rangle\langle\gamma_k|$$
$$= \sum_k |\langle\alpha_k,\gamma\rangle|^2 P_{\gamma_k}.$$

$$Tr_1\rho_M = \sum_n\langle\alpha_n|\sum_k|\langle\alpha_k,\gamma\rangle|^2 P_{\alpha_k}\otimes P_{\gamma_k}|\alpha_n\rangle$$
$$= \sum_k |\langle\alpha_k,\gamma\rangle|^2\left(\sum_n\langle\alpha_n,P_{\alpha_k}\alpha_n\rangle\right)P_{\gamma_k}$$
$$= \sum_k |\langle\alpha_k,\gamma\rangle|^2 P_{\gamma_k}$$

$$Tr_2\rho = \iiint dx\,dy\,dz\,\langle\delta_{xyz}|\sum_\ell\langle\alpha_\ell,\gamma\rangle\alpha_\ell\otimes\gamma_\ell\rangle\langle\sum_k\langle\alpha_k,\gamma\rangle\alpha_k\otimes\gamma_k|\delta_{xyz}\rangle$$
$$= \sum_{k\ell}\iiint dx\,dy\,dz\,\langle\alpha_\ell,\gamma\rangle\langle\gamma,\alpha_k\rangle|\alpha_\ell\rangle\langle\alpha_k|\langle\delta_{xyz},\gamma_\ell\rangle\langle\gamma_k,\delta_{xyz}\rangle$$
$$\cong \sum_k |\langle\alpha_k,\gamma\rangle|^2 P_{\alpha_k}, \text{ since } \langle\delta_{xyz},\gamma_\ell\rangle, \langle\delta_{xyz},\gamma_k\rangle$$

$k\neq\ell$, are each nonzero in disjoint X, Y, Z-regions.

$$Tr_2\rho_M = \iiint dx\,dy\,dz\,\langle\delta_{xyz}|\sum_k|\langle\alpha_k,\gamma\rangle|^2 P_{\alpha_k}\otimes P_{\gamma_k}|\delta_{xyz}\rangle$$
$$= \sum_k |\langle\alpha_k,\gamma\rangle|^2\left(\iiint dx\,dy\,dz\,\langle\delta_{xyz},P_{\gamma_k}\delta_{xyz}\rangle\right)P_{\alpha_k}$$
$$= \sum_k |\langle\alpha_k,\gamma\rangle|^2 P_{\alpha_k}.$$

It is therefore tempting to declare that the total ensemble prepared by a Stern-Gerlach device may be split into pure subensembles with state vectors $\alpha_k\otimes\gamma_k$, $k=1,2$, provided only observables in $\mathcal{A}$ are considered.

---

*This equivalence class is not complete; i.e., there do exist operators outside $\mathcal{A}$ for which the two ρ's are equivalent. Gottfried[74], e.g., uses in the Stern-Gerlach problem a class different from our $\mathcal{A}$.

Hence selection of subensemble $P_{\alpha_1\otimes\gamma_1}$ would constitute $\Pi(P_{\alpha_1\otimes\gamma_1};\mathcal{A})$ i.e., preparation of the state $\alpha_1\otimes\gamma_1$ relative* to $\mathcal{A}$. If this conclusion were correct, then our above critique of the "filtration" argument would be reduced in practical cases almost to verbal quibbling; however, owing to an inherent weakness of the "relative preparation" concept $\Pi(P_{\alpha_1\otimes\gamma_1};\mathcal{A})$, we still insist that the "filtration" argument does not justify the claim that the apparatus in question-- $\mathcal{H}\otimes\mathcal{K}_0$ -source, Stern-Gerlach magnet, and W --prepares $P_{\alpha_1\otimes\gamma_1}$ in any sense. The trouble with "relative preparation" $\Pi(\rho;\mathcal{A})$ is its reliance upon $\mathcal{A}$-equivalence, which is not a temporally invariant property. Although it is true, as demonstrated above, that $Tr(\rho A)=Tr(\rho_M A)$, for each A in $\mathcal{A}$ at a given instant, say t_1, it does not follow that $Tr(T\rho T^{\dagger}A)=Tr(T\rho_M T^{\dagger}A)$ for each A in $\mathcal{A}$, where $T\equiv T(t_2,t_1)$. Consider, for example, the probability densities $w(\alpha_n,X,Y,Z;\rho(t))$ and $w(\alpha_n,X,Y,Z;\rho_M(t))$

$$w(\alpha_n,X,Y,Z;\rho(t))=Tr(\rho(t)P_{\alpha_n\otimes\delta_{XYZ}})=\left|\sum_{\beta\gamma}\langle\alpha_\beta,\gamma\rangle\langle\alpha_n\otimes\delta_{XYZ},T(\alpha_\beta\otimes\gamma_\gamma)\rangle\right|^2,$$

$$w(\alpha_n,X,Y,Z;\rho_M(t))=Tr(\rho_M(t)P_{\alpha_n\otimes\delta_{XYZ}})=\sum_{\beta\gamma}\left|\langle\alpha_\beta,\gamma\rangle\langle\alpha_n\otimes\delta_{XYZ},T(\alpha_\beta\otimes\gamma_\gamma)\rangle\right|^2.$$

At $t=t_1$, $T=1$, and these expressions are of course equal; but for $t>t_1$, $T\neq 1$ and the distributions are unequal. This demonstrates that the $\mathcal{A}$-equivalence of ρ and ρ_M at t_1 is useful only in statics; in general, even if the only observables ever considered are in $\mathcal{A}$, still ρ cannot be replaced by ρ_M for dynamic applications. Only ρ correctly represents the idea of state in its causal role; ρ_M, while equivalent to ρ at t_1, leads to incorrect predictions and cannot therefore be regarded as the state. Hence the foregoing "filtration" argument cannot be accepted as a quantal explanation of the state preparation $\Pi(P_{\alpha_1\otimes\gamma_1})$.

Thus far, it has been left undecided whether or not the combination of $\mathcal{H}\otimes\mathcal{K}_0$ -source, Stern-Gerlach magnet, and W may actually be used to

*This should not be confused with Everett's concept of "relative state".[80]

prepare $P_{\alpha_1 \otimes \gamma_1}$; all that has been established is the classical naivete of the common description in terms of "filtration". The success of failure of the proposed device as a $P_{\alpha_1 \otimes \gamma_1}$-preparer depends mainly upon the nature of the system W. The $\psi \otimes \chi_0$-source plus Stern-Gerlach magnet has prepared $P_{\sum_k \langle \alpha_k, \psi \rangle \alpha_k \otimes \gamma_k}$; what must be shown is that the interaction with W could convert this pure ensemble into a mixture from which the pure subensemble $P_{\alpha_1 \otimes \gamma_1}$ might be extracted. We shall now discuss two model theories of preparation which would explain in a sensible quantum theoretic way how the combined apparatus in question could perform $T(P_{\alpha_1 \otimes \gamma_1})$.

(1) For a simple but rather fanciful model, assume W is a slab of antimatter initially containing N antiatoms. Since annihilation processes connected with matter-antimatter interaction must be considered, we need three Hilbert spaces: $\mathcal{H}_S$, associated with the atom; $\mathcal{H}_W$, with the slab; and $\mathcal{H}_F$, with electromagnetic radiation. Let ϕ_0 denote the radiation vacuum state, ϕ some other radiation state, ω_N a stationary state for the N antiatoms, γ_{N-1} some state for (N-1) antiatoms, and Ψ_0 the atom's vacuum state. Suppose the initial state of the composite system $S + W + F$ is $\left(\sum_k \langle \alpha_k, \psi \rangle \alpha_k \otimes \gamma_k\right) \otimes \omega_N \otimes \phi_0$. If U denotes the evolution operator for this total system from the preparation of the above initial state until a time when the atomic wave packet would have passed beyond the absorber position if W were absent, then in terms of U-matrix elements, the following assumptions define a W capable of producing the desired mixtures:

$$\langle K \otimes \lambda \otimes \xi \mid U \mid (\alpha_1 \otimes \gamma_1) \otimes \omega_N \otimes \phi_0 \rangle = 0$$

unless $K = T(\alpha_1 \otimes \gamma_1)$, $\lambda = \omega_N$, $\xi = \phi_0$, where T describes the evolution of the atom when W is absent.

$$\langle K \otimes \lambda \otimes \xi \mid U \mid (\alpha_2 \otimes \gamma_2) \otimes \omega_N \otimes \phi_0 \rangle = 0,$$

unless $K=\Psi_0, \lambda=\chi_{N-1}$. (Classically speaking, these expressions mean that an atom may traverse region 1 undisturbed but in region 2 would be destroyed.)

The final state of the composite system would therefore be of the form

$$U\left[\left(\sum_k \langle \alpha_k, \psi\rangle \alpha_k \otimes \gamma_k\right) \otimes \omega_N \otimes \phi_0\right]$$
$$= \langle \alpha_1, \psi\rangle [T(\alpha_1 \otimes \gamma_1)] \otimes \omega_N \otimes \phi_0 + \langle \alpha_2, \psi\rangle \Psi_0 \otimes \chi_{N-1} \otimes \phi.$$

"Tracing out" the antimatter and radiation parts, we find* that the atom-ensemble now has the density operator $\rho_1 = |\langle \alpha_1, \psi\rangle|^2 P_{T(\alpha_1 \otimes \gamma_1)} + |\langle \alpha_2, \psi\rangle|^2 P_{\Psi_0}$, which shows that only the fraction $|\langle \alpha_1, \psi\rangle|^2$ of the atoms from the original $\psi \otimes \chi_0$-source emerge from the complete apparatus and that this sub-ensemble has state vector $T(\alpha_1 \otimes \gamma_1)$. (For the short time interval of interest, $T \cong 1$; hence we have effectively a $\Pi(P_{\alpha_1 \otimes \gamma_1})$-device.)

(2) A somewhat more realistic model results if W is taken as a slab of ordinary matter. If U again denotes the evolution operator during interaction, the idea of "region 2 absorber" may perhaps be expressed as follows: $U[(\alpha_1 \otimes \gamma_1) \otimes \omega_N] = T(\alpha_1 \otimes \gamma_1) \otimes \omega_N \cong (\alpha_1 \otimes \gamma_1) \otimes \omega_N$,

$$U[(\alpha_2 \otimes \gamma_2) \otimes \omega_N] = \sum_{\ell m} \nu_{\ell m} \phi_\ell \otimes \omega_m$$

, where $\{\omega_k\}$ is a complete set of stationary states for W, $\{\phi_\ell\}$ is a complete eigenvector set associated with the atom, and $\nu_{\ell m}$ has these properties: $\nu_{\ell N} \cong 0$, for each ℓ; $\nu_{\ell m} \cong 0$, unless $\sum_n |\langle \alpha_n \otimes \delta_{xyz}, \phi_\ell\rangle|^2$ is essentially nonzero only within the slab, for each m.

The final state of the total system is then

$$U\left[\left(\sum_k \langle \alpha_k, \psi\rangle \alpha_k \otimes \gamma_k\right) \otimes \omega_N\right]$$
$$\cong \langle \alpha_1, \psi\rangle (\alpha_1 \otimes \gamma_1) \otimes \omega_N + \langle \alpha_2, \psi\rangle \sum_{\ell m} \nu_{\ell m} \phi_\ell \otimes \omega_m.$$

Let $\sum_\ell \nu_{\ell m} \phi_\ell = g_m \theta_m$, $\langle \theta_m, \theta_m\rangle = 1$, $m \neq N$;

$\langle \alpha_1, \psi\rangle = g_N$, $\alpha_1 \otimes \gamma_1 = \theta_N$; and $\langle \alpha_2, \psi\rangle g_m = g_m$, $m \neq N$.

*This tracing is easily done if one just notes that $\langle T(\alpha_1 \otimes \gamma_1), \Psi_0\rangle = \langle \omega_N, \chi_{N-1}\rangle = \langle \phi_0, \phi\rangle = 0$; the general mathematical form is then the same as that encountered in the correlation assumption of orthodox measurement theory (section 6).

With these substitutions, the final state becomes $\sum_k q_k \theta_k \otimes w_k$ where $\{w_k\}$ is an orthogonal set but $\{\theta_k\}$ is not. Now, consider the reduced density operator for the atom alone:

$$\begin{aligned}\rho_1 &= \mathrm{Tr}_W P_{\sum_k q_k \theta_k \otimes w_k} \\ &= \sum_m \langle w_m | \sum_\ell q_\ell \theta_\ell \otimes w_\ell \rangle \langle \sum_k q_k \theta_k \otimes w_k | w_m \rangle \\ &= \sum_{k\ell m} q_\ell q_k^* |\theta_\ell\rangle\langle\theta_k| \delta_{\ell m} \delta_{km} \\ &= \sum_k |q_k|^2 P_{\theta_k} = |\langle \alpha_1, \psi \rangle|^2 P_{\alpha_1 \otimes \gamma_1} + \sum_{k \neq N} |q_k|^2 P_{\theta_k}.\end{aligned}$$

Thus the atom ensemble has become a genuine mixture from which it is possible to select the subensemble $P_{\alpha_1 \otimes \gamma_1}$. Since all other subensembles are localized in the slab by the interaction, we may conclude that the ensemble of atoms emerging from the complete apparatus (the fraction $|\langle \alpha_1, \psi \rangle|^2$ of the original atoms) has state vector $\alpha_1 \otimes \gamma_1$. This model therefore exemplifies a rational quantum mechanical explanation of a $\Pi(P_{\alpha_1 \otimes \gamma_1})$-device; yet no measurement$_2$ process was involved.

It should now be clear that $\Pi(P_{\alpha_1 \otimes \gamma_1})$ is a physical process different from $M_2(S_z)$ but inexplicable without tacit reference to M_1's. In particular, there is no justification for any general statement that $\Pi(P_{\alpha_1 \otimes \gamma_1})$ is equivalent to an $M_2(S_z)$ yielding α_1. The apparatus just described--Stern-Gerlach magnet plus region 2 absorber--would effect $\Pi(P_{\alpha_1 \otimes \gamma_1})$; but it would not perform $M_2(S_z)$ (unless the detection of <u>nothing</u> in region 1 is regarded as a measurement of S_z yielding α_1 !). Indeed the <u>preparation</u> of the $P_{\alpha_1 \otimes \gamma_1}$-ensemble still occurs even if the absorber is not a detector at all, i.e., even if it simply does not record whether or not it captured any atom.

The Stern-Gerlach example of the past few sections has demonstrated that the constructs M_1, M_2 and Π should be carefully distinguished in "measurement theories"; otherwise ambiguity and confusion are inevitable. An excellent example of this confusion is the term "selective measurement";[81]

which is sometimes used quite indiscriminately to refer to both measurement and preparation whenever a "separation and filtration" scheme like the Stern-Gerlach device is the physical basis of both Π and $\mathcal{M}_2$. As a result, it is virtually impossible to determine the meaning or purpose of a discussion on so-called "selective measurements". The distinction between $\mathcal{M}_1, \mathcal{M}_2$, and Π is abnormally subtle in the Stern-Gerlach, or "selective", case; for this reason, it is a favorite example among proponents of wave packet reduction and/or the equivalence of $\mathcal{M}_2(a)$ and $\Pi(P_{\psi_R})$. Nevertheless, as we have seen, the differences among these concepts can still be exposed.

Ordinary applications of quantum theory are normally successful in spite of the occasional confusion of preparation and measurement. However, the distinction can be quite important in basic theoretical considerations. A good example of faulty reasoning due to the implicit assumption that preparation must be accomplished through measurement occurs in connection with the superselection rules[2] of quantum field theory. These rules arise from invariance principles which, applied to states, require that certain distinct state vectors (rays) be physically equivalent. Now, there are at least two ways to guarantee this equivalence: (1) Postulate that not all Hilbert vectors represent physically realizable states, or equivalently that there is no process $\Pi(P_\psi)$ for certain ψ's. The Pauli principle is a familiar example of such a requirement. In the case of superselection rules, it turns out that the distinct state vectors which must be equivalent can be eliminated by postulating that all physically realizable state vectors are eigenvectors of certain operators (total charge, e.g.). (2) Modify the common axiom that all Hermitean operators represent observables by explicitly denying the "observability" of all operators having different mean values for those state vectors which must

be equivalent.*

The concept of superselection rule is relatively new and still under development. It is therefore impossible to make very definite statements about it.** The only point to be made here is that a standard "theorem" which purports to derive (2) from (1) is fallacious because it confuses measurement and preparation. A typical presentation of the argument is given by Schweber:[82] "If not all rays are realizable, then clearly no measurement can give rise to these nonrealizable states. They cannot therefore be eigenfunctions of any Hermitean operator which corresponds to an observable property of the system. To be observable a Hermitean operator must therefore satisfy certain conditions (superselection rules)." The first sentence is incontestable; indeed, if a state vector cannot be prepared at all, certainly no measurement process can do the job. The second sentence is a non sequitur obviously based on the false premise that an $M_2(a)$ yielding a_R is the same as $\Pi(P_{a_R})$. Thus if $\Pi(P_{a_R})$ is impossible, $M_2(a)$ must be impossible. The third sentence follows from the second and is just alternative (2) above. We see therefore that confusion of measurement and preparation is here responsible for the theoretical illusion that (2) is a consequence of (1), whereas in fact (2), if needed, should be postulated independently.

Although M_1, M_2, and Π must never be regarded as equivalent, there are of course connections among them which we do not wish to deny.

---

*Cf. fn., p. 3.

**In our opinion alternative (1), a natural generalization of the Pauli principle, is preferable to (2) and is all that is really needed to account for physical facts of the type which suggest the existence of superselection rules, e.g., the fact that the superposition of an electron state vector and a positron state vector apparently describes no actual ensemble. If (2) were unnecessary, the word some would be unnecessary in P1, and P2 would not have to make the rather odd demand that $m(A)$ be real even when A is a Hermitean operator representing no observable.

These relations may be expressed as follows: (1) M_2 and Π are both laboratory processes the quantum theoretical description of which necessarily involves the primitive construct M_1. (2) $M_2(a)$, like any physical process, leaves the systems involved in some state; and the ensembles of these systems would have calculable density operators. In this trivial sense, all physical processes, $M_2(a)$ included, prepare states. However, these states need not exhibit any special relation to a; in particular, preparations effected by $M_2(a)$ are not necessarily, nor even usually, $\Pi(P_{a_k})$. To defend his use of an a_k in a complex situation where $M_2(a)$ would never be performed, Schrödinger once remarked: "A purist might challenge the use of a wave function not determined by measurement. But he would have to give up using wave functions altogether, since none has ever been determined by measurement."[83] (3) Similarly, the physical process $\Pi(P_{a_k})$ might conceivably be utilized as an M_2, perhaps even $M_2(a)$; but this need not be the case. Usually all that is known from $\Pi(P_{a_k})$ is that, if $M_2(a)$ were performed, the correlations thereby established would show that an a-measurement, i.e., $M_1(a)$, must yield a_k.

20. Summary: Quantum Theory of Measurement

The problem of quantum measurement was introduced in section 1 in the customary way as a logical challenge to be met within the quantal framework. At issue was the fact that the explicit appearance of the term measurement in the postulates of quantum theory automatically confers some properties upon that concept. Yet the notion of measurement does not really belong to quantum physics in particular; indeed measurement is basic to all of physical science and presumably comes to quantum theory already endowed with characteristics inherent in its more general epistemological role. A "quantum theory of measurement" would then be a confrontation

of the measurement concept in quantum theory with the idea of measurement in general in order to demonstrate the consistency of the quantum viewpoint.

However, a fundamental defect in this program gradually became apparent. The quantum measurement construct is well defined in the sense that the postulates offer clear instructions as to its use in theoretical explanations of physical processes. On the other hand, the general philosophical understanding of measurement cannot be expressed in simple mathematical terms. To demonstrate this point, we critically surveyed the major classes of proposed quantum measurement theories; invariably, the extraquantal strictures placed upon measurement were found to result in physically unwarranted "overspecifications" of its root meaning. Scrutiny of these overly narrow definitions of measurement served mainly to expose misunderstandings about the nature of quantum theory.

Eventually we recognized that this entire approach was foredoomed. Even if a grand, all-embracive mathematical definition of a general measurement process were discovered, it could not serve to establish the consistency of the quantum theoretical usage of the term measurement; for no physical process, measurement schemes included, can be described by quantum theory without the term measurement, which has therefore the logical status of an ultimate primitive, irreducible to other constructs.

This state of affairs, essentially a consequence of the latency of quantum observables, suggested the necessity of distinguishing between $\mathcal{M}_1$ and $\mathcal{M}_2$. The nature of these two measurement constructs as well as their significance for quantum measurement theory was explained in several preceding sections. However, to clarify these ideas, we shall now briefly recapitulate by describing $\mathcal{M}_1$ and $\mathcal{M}_2$ in another way, viz., by focusing upon their epistemological status. To insure direct passage through the sometimes labyrinthine halls of scientific epistemology, it is helpful to

refer to a chart (Fig. 1) originated by Margenau.[84]

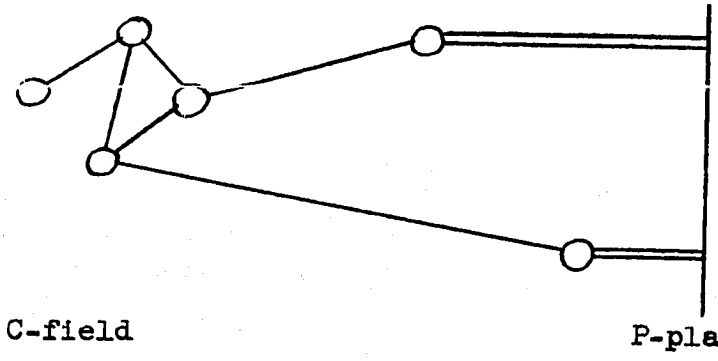


Fig. 1

A simplified legend for this "epistemological map" would make these identifications: (1) The P-plane represents uninterpreted sense impressions, those elements of experience variously called the given, the percepts, the direct observations, data, and by Margenau, the protocols. (2) The C-field is the domain of reason, of ideal models; its members (denoted by circles) are known as the categories, concepts, or constructs; and in Einstein's words, they are "free creations of the human mind".[85] (3) A set of rational connections (single lines) among constructs forms the logical matrix of a theory. (4) Some constructs are related to direct observations at the perceptual level (P-plane) by conventions (double lines) which may be called operational definitions (Bridgman[86]), rules of correspondence (Margenau[87]), or epistemic correlations (Northrop[88]). (5) The distance of a construct from the P-plane is to be regarded as an indication of its relative abstractness, or, in a sense, its objectivity. This horizontal "scale" is of course rather vague, but it is not meaningless. For example, the construct "electric field" is obviously far to the left of "electric shock"; the sequence of concepts "entropic-derivative-of-internal-energy", "thermometric-temperature", and "hotness" evidently range from extremely far into the C-field to extremely close to the P-plane.

The general notion of measurement as a universal feature of the

scientific method--what we have designated $\mathcal{M}_2$--refers to an important part of the complex of linkages between the most profound constructs and the practically self-evident ones just short of the diffuse boundary of raw, undifferentiated percepts. Measurement$_2$ is concerned directly with those constructs called observables which mediate between mathematical models and direct observations; the defining characteristic of any measurement$_2$ scheme is therefore the extraction of numbers from observations and their theoretically meaningful assignment to the observables. The overall purpose and pragmatic value of this procedure is fully discussed in books on the philosophy of science, but these matters are not at issue here.

With the above understanding of measurement$_2$ as the provider of numbers to observables, it is easy to describe in a general way, using Margenau charts, just what a "theory of measurement " would be. Consider an observable $\mathcal{a}$ which is defined constitutively by the properties of its representative A among the mathematical constructs of the theory. Suppose that an operation has been discovered the performance of which yields numbers and that these numbers can be consistently associated with $\mathcal{a}$, in some sense, as its "values". The measurement concept $\mathcal{M}_2(\mathcal{a})$ is then simply the rule of correspondence which specifies that operation (Fig. 2).

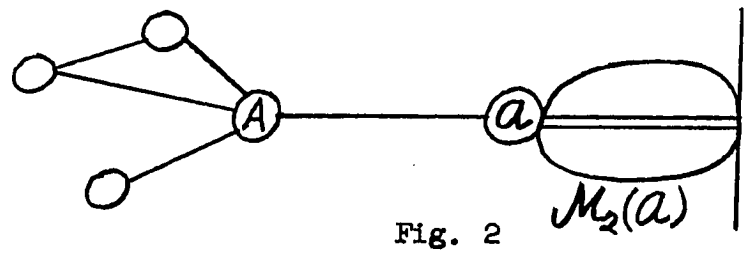


Fig. 2

A theory of measurement$_2$ is then simply a theoretical analysis of the operation identified as $\mathcal{M}_2(\mathcal{a})$. In other words, part of the rule of correspondence itself is explained in terms of fundamental constructs. As a result, the concept $\mathcal{M}_2(\mathcal{a})$ acquires a more complex "structure" (Fig. 3).

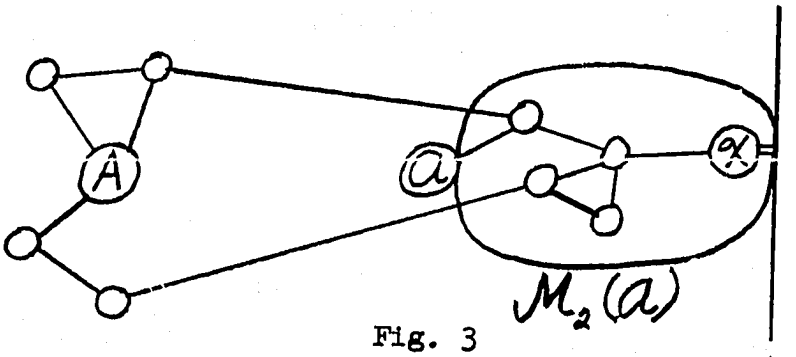


Fig. 3

On the epistemological chart, the unanalyzed double line connecting a to the distant P-plane is replaced by theoretical connections from a to an observable x plus a very short double line from x to the near-by protocols. The closeness of x to the P-plane signifies that, as far as physics is concerned, x is regarded as directly observable. In earlier sections, x was taken to be position, said to be measured by "looking at" the coincidence of a "spot" and a "scale marking". However, we adopted this specific identification of x, suggested by the writings of deBroglie and Landé, only to exemplify the ultimate contact of physical theory and empirical experience; there is no reason to regard it as the sole direct observable of potential value for physics.

What has been said about measurement$_2$ thus far has been applicable to science in general. Hence problems motivated by the foregoing remarks cannot be legitimately interpreted as quantum dilemmas in particular. For example, the unanalyzed connection of an x to the diffuse realm of immediacy suggests the problem of infinite regression quite independently of quantum theory, as noted in section 15. At any rate, the measurement concept $M_2(a)$ is epistemologically the same in quantum physics as in the rest of science; and a quantum theory of $M_2(a)$ should be of no more philosophic interest than are classical disciplines such as thermometry and photometry.

On the other hand, the measurement concept M_1 is peculiar to quantum theory. To find its proper "location" on the Margenau chart,

recall that the essence of measurement$_2$ is the extraction of numbers from observations and their theoretically meaningful assignment to the observables. Hence the description of an $M_2(a)$ must always be given in terms of such assignments; but we have seen in previous sections that classical and quantal physics do not employ the same relation between an observable and its numerical values. Classically, an observable is said to "have" its value; quantally, the only connection is through the auxiliary measurement concept $M_1(a)$. Accordingly, on the epistemological chart, let us replace the observable symbol a by $y(a)$ if a "has" a value and by $M_1(a)$ if the latent results of potential $M_1(a)$'s represent the only connection between and its values. Comparison of the classical (Fig. 4) and quantal (Fig. 5) realizations of Fig. 3 then serves to clarify the epistemological status of M_1.

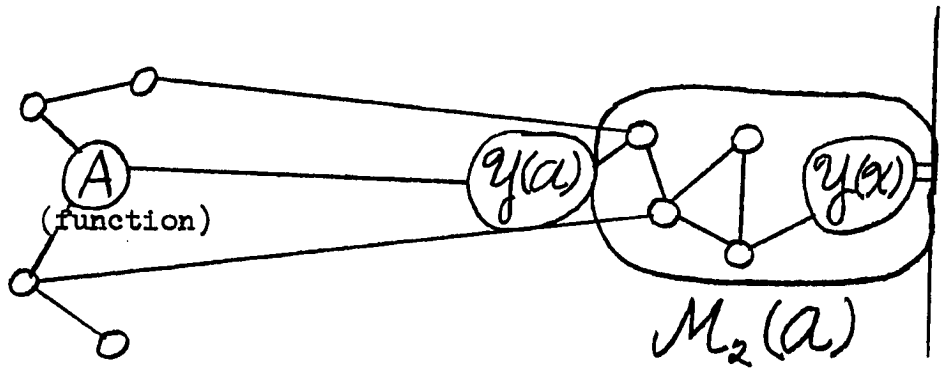


Fig. 4

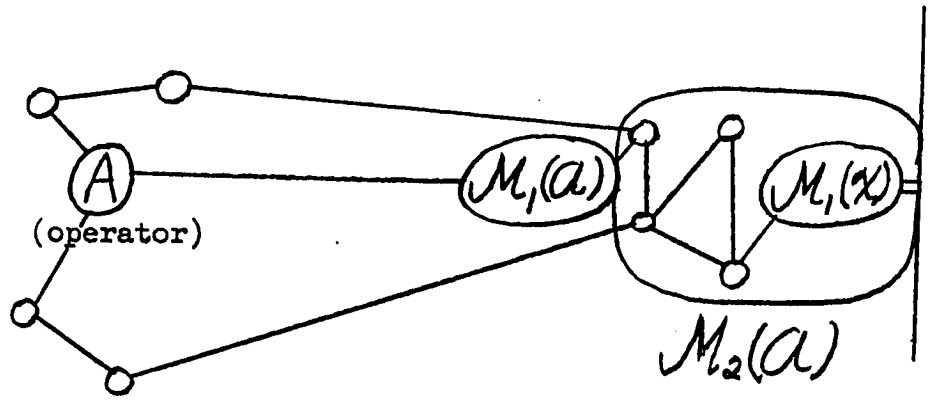


Fig. 5

These charts represent an attempt to summarize graphically the main points about quantum measurement concepts discussed in previous sections. In particular, they emphasize that $\mathcal{M}_1$ is analogous to the classical idea y_0 of "having" a value. It is meaningless to consider further analysis of either y or $\mathcal{M}_1$; logically, both are ultimate primitives in their respective theories in the sense that no physical process can be described without them. There could no more be a "quantum theory of $\mathcal{M}_1$" than there could be a "classical theory of y"; either would be quite circular. Hence the term "quantum theory of measurement" can only refer to a theory of $\mathcal{M}_2(a)$, the statement of which will necessarily employ $\mathcal{M}_1$'s. Understood in this way, so-called quantum theories of measurement$_2$ are of no more of less philosophic interest than analogous classical theories of measurement$_2$ which explain the operation of calorimeters, spectrometers, etc. Indeed the rather extraordinary qualities sometimes attributed to quantum measurement derive from the various misinterpretations of quantum theory which the present work has sought to expose. Once the distinctness of $\mathcal{M}_1$ and $\mathcal{M}_2$ is recognized, the general concept of measurement $\mathcal{M}_2$ is no more mysterious in quantum physics than elsewhere.

Selected References

1. E.g.: D. W. Mackey, The Mathematical Foundations of Quantum Mechanics, Benjamin (1963); M. D. Maclaren, "Notes on Axioms for QM", Argonne Nat'l Lab. ANL-7065.

2. G. C. Wick, A. S. Wightman, and E. P. Wigner, Phys. Rev., 88 (1952), 101.

3. J. L. Park, "Simultaneous Measurement in Quantum Theory" (Part III of Thesis).

4. J. L. Park, "The Nature of Quantum States" (Part I of Thesis).

5. J. von Neumann, Mathematical Foundations of Quantum Mechanics, Eng. trans. by R. T. Beyer, Princeton (1955), 313-316; cf. also A. M. Gleason, J. Math. Mech., 6 (1957), 885.

6. von Neumann, op. cit., 295-313.

7. Ref. 6, p. 323.

8. H. Margenau, Philosophy of Science, 4 (1937), 352-356; 25 (1958), 23.

9. E.g.: N. Bohr, Essays 1958-1962 on Atomic Physics and Human Knowledge, Interscience (1963), 1-7.

10. H. Margenau, The Nature of Physical Reality, McGraw-Hill (1950), 171-177.

11. W. Heisenberg, Physics and Philosophy, Harper (1958), 53.

12. Ref. 10, esp. Ch. 13.

13. R. Carnap, Philosophy and Phenomenological Research, 5 (1945).

14. D. ter Haar, Elements of Statistical Mechanics, Rinehart (1954), 149.

15. H. Margenau, Philosophy of Science, 30 (1963), 138.

16. Ref. 6, p. 332.

17. Ref. 6, p. 321; cf. also J. M. Jauch and G. Piron, Helvetica Physica Acta, 36 (1963), 827-837.

18. E.g.: Ref. 6, pp. 437-445; F. London and E. Bauer, Actualities Scientifiques et Industrielles No. 775, Hermann et Cie (1939), 22-47; Ref. 15, pp. 38-41.

19. Ref. 6, pp. 422-437

20. Ref. 6, p. 347.

21. E.g.: D. Bohm, Quantum Theory, Prentice-Hall (1951), 120-129.

22. C. J. Davisson, Franklin Institute Journal, 205 (1928), 597.

23. W. Heisenberg, The Physical Principles of Quantum Theory, Eng. trans. by C. Eckart and F. C. Hoyt, Chicago (1930), 59.

24. Ibid., 61.

25. Ibid., 43-46.

26. W. Heisenberg in Neils Bohr and the Development of Physics, W. Pauli (Ed.), Pergamon (1955), 25.

27. Ref. 26, p. 26.

28. N. Bohr, J. Chem. Soc., 135 I (1932), 376-377.

29. Ref. 11, p. 53.

30. Ref. 26, p. 27.

31. E.g.: G. Süssmann in Observation and Interpretation, S. Korner (Ed.), Butterworth (1957), 131-136; J. S. Bell and M. Nauenberg in Preludes in Theoretical Physics, de-Shalit et al. (Ed.), North-Holland (1966), 279-285; References 53, 55, 56, 57, 80, and many others.

32. D. I. Blochintsev, Quantum Mechanics, D. Reidel (1964), 494-501.

33. Ref. 11, p. 138.

34. Ref. 32, pp. 55-59.

35. E.g.: F. Mandl, Quantum Mechanics, Butterworth (1954), 68-69.

36. W. Heitler in Albert Einstein: Philosopher-Scientist, W. Schilpp (Ed.), Tudor (1949), 190.

37. N. F. Mott, Proc. Roy. Soc., 126 (1929), 79.

38. L. D. Landau and Lifshitz, Quantum Mechanics, Pergamon (1958), 26-27.

39. P. Ehrenfest, Zeit. fur Phys. 45 (1927), 455.

40. Ref. 3, sec. 9.

41. E. P. Wigner, Zeitschrift fur Physik, 133 (1952), 101.

42. H. Araki and M. Yanase, Phys. Rev., 120 (1961), 622.

43. M. Yanase, Am. J. of Phys., 32 (1964), 208.

44. W. Pauli in Handbuch der Physik, 24, Springer (1933).

45. E.g.: P. K. Feyerabend in Frontiers of Science and Philosophy, R. G. Colodny (Ed.), Pittsburgh (1962), 200.

46. Ref. 38, pp. 21-24.

47. L. Durand, III, Philosophy of Science, 27 (1960), 115.

48. E. Schrödinger, Naturwissenschaften, 23 (1935), 812.

49. E.g.: Discussion, pp. 138-147, in Observation and Interpretation, S. Korner (Ed.), Butterworth (1957).

50. Ref. 38, p. 22.

51. Ref. 38, p. 2.

52. Ref. 6, pp. 437-438.

53. E. P. Wigner, Am. J. Phys., 31 (1963), 11-12.

54. A. Komar, Phys. Rev., 126 (1962), 365.

55. P. K. Feyerabend in Observation and Interpretation, S. Korner (Ed.), Butterworth (1957), 121-130.

56. H. Wakita, Progress of Theoretical Physics, 23 (1960), 32.

57. J. M. Jauch, Helvetica Physica Acta, 37 (1964), 293.

58. N. Bohr in Albert Einstein: Philosopher-Scientist, P. Shilpp (Ed.), Tudor (1949), 209.

59. Ref. 11, p. 44.

60. Ref. 21, pp. 624-628.

61. Ref. 45, p. 264.

62. N. R. Hanson, The Concept of the Positron, Cambridge (1963), 60-70.

63. Ref. 38, p. 3.

64. G. Bergmann, Am. J. Phys., 15 (1947).

65. P. Jordan, Philosophy of Science, 16 (1949), 269.

66. G. Ludwig, Die Grundlagen der Quantenmechanik, Springer (1954), 122-165.

67. Ref. 53, p. 14.

68. A. Daneri, A. Loinger, and G. M. Prosperi, Nuclear Physics, 33 (1962), 297; Nuovo Cimento B 44 (1966), 119.

69. L. Rosenfeld, Supplement of the Progress of Theoretical Physics, Extra number (1965), 229.

70. J. B. Stallo, The Concepts and Theories of Modern Physics, Appleton (1881) and Harvard (1960).

71. Ref. 10; A. Landé, New Foundations of Quantum Mechanics, Cambridge (1965).

72. W. Gerlach and O. Stern, Zeitschrift fur Physik, 9 (1922), 349, translated by H. A. Boorse and L. Motz (Eds.) in The World of the Atom, Basic Books (1966), 936.

73. T. E. Phipps and J. B. Taylor, Phys. Rev., 29 (1927), 309.

74. E.g.: Ref. 21, pp. 593-602; K. Gottfried, Quantum Mechanics, Benjamin (1966), 165-189.

75. Landé, op. cit., 121.

76. L. deBroglie, The Current Interpretation of Wave Mechanics, Elsevier (1964), 4.

77. J. Schwinger, Proc. Nat'l. Acad. Sci. U. S., 46 (1960), 570.

78. H. J. Groenewold, Nuclear Physics, 57 (1964), 133.

79. Ref. 26, p. 27; Ref. 75, pp. 128-131.

80. H. Everett III, Rev. Mod. Phys., 29 (1950), 454.

81. Ref. 77; F. A. Kaempffer, Concepts in Quantum Mechanics, Academic press (1965), 31-38.

82. S. S. Schweber, An Introduction to Relativistic Quantum Field Theory, Row, Peterson (1961), 6.

83. E. Schrödinger, Statistical Thermodynamics, Cambridge (1952), 90.

84. Ref. 10, pp. 102-107; Ref. 15, pp. 1-2.

85. A. Einstein, "Geometry and Experience", Readings in the Philosophy of Science, H. Feigl and M. Brodbeck (Eds.), Appleton-Century-Crofts (1953), 189.

86. P. W. Bridgman, The Logic of Modern Physics, Macmillan (1928).

87. H. Margenau, Phil. of Sci. 2 (1935), 1.

88. F. S. C. Northrop, The Logic of the Sciences and the Humanities, Macmillan (1947).

III. SIMULTANEOUS MEASURABILITY IN QUANTUM THEORY

ABSTRACT

Part III is a study of what is sometimes regarded as the conceptual heart of quantum theory, viz., the orthodox "physical" interpretation of noncommuting operators as representatives of incompatible (non-simultaneously-measurable) observables. To provide a firm foundation for the analysis, a definite statement of the essentials of modern quantum theory is given briefly in the form of a mathematical axiomatization together with a review of the two measurement constructs introduced in Part II. Contrary to custom in discussions on simultaneous measurability, the uncertainty principle is not dwelt upon but simply stated carefully to establish its actual irrelevance to the problem at hand. It is then demonstrated that the much quoted "principle" of incompatibility of noncommuting observables is simply false. The axiomatic root of all incompatibility arguments is next identified; and it is shown that, with a slight modification of the basic postulates which affects neither useful theorems nor practical calculations, quantum physics no longer entails illogical restrictions on measurability. Among the related topics touched upon are the problem of joint probability distributions, the "logical" approach to quantum mathematics (wherein noncommutativity becomes incompatibility within a propositional calculus), and the field theoretic attempt to unify quantal and relativistic physics through a postulated connection between incompatibility and space-like intervals.

1. The Compatibility Problem

It is characteristic of physics to represent observables by mathematical objects to which the numbers emerging from experiments are to be related. In classical physics the objects were functions, and numerical measurement results were merely identified with the range values of these functions. Quantum physics, on the other hand, uses (Hermitean) operators instead of functions, and thereby complicates the relationship between its observables (i.e., their mathematical representatives) and the empirical numbers to which they must ultimately refer. Ironically, perhaps the most abstruse and controversial difficulty associated with such operator-observable correspondence arises from an obvious arithmetical law, viz., that if a, b are numbers, $ab = ba$. Naturally this commutative law applies to all measurement results independently of the theory by which they are interpreted. In classical physics, this numerical commutativity is reflected by the unrestricted commutativity of functions; in quantum theory, however, the analogous statement cannot be made, for pairs of Hermitean operators do not necessarily commute. In particular, it was discovered by Born in the early years of modern quantum mechanics that even X and P, the operators representing the important observables position and momentum, do not commute but obey instead his famous equation $[X,P] \equiv XP - PX = i\hbar 1$. Thus position and momentum are said to be noncommuting observables.

Understandably, the presence in quantum theory of noncommuting observables has from the beginning elicited a great deal of academic curiosity accompanied by the reasonable suspicion that such a theoretical anomaly cannot merely be written off as a mathematical quirk. Some kind of physical interpretation must be given; the fact that $[X,P] \neq 0$ surely expresses something very interesting about position and momentum. But what? The orthodox answer is this: noncommuting observables are incompatible, i.e.,

it is impossible to perform upon a single system simultaneous measurements of two such observables. The present work is devoted to the systematic analysis of this famous principle of impotence; but first, as a prelude to this endeavor, it seems appropriate to review briefly the more common--and frequently illogical--arguments typically advanced in behalf of the doctrine in question.

(1) Semiclassical gedankenexperiments: It is quite fashionable in discussions on the foundations of quantum theory to lean heavily on the historical evolution of the subject. This tendency is not new to physics; it is in fact traditionally employed in studies of relativity, thermodynamics, and electrodynamics. But while the origin of any of these disciplines constitutes a fascinating chapter in the history of physical ideas, the relevance of chronological development to logical development must not be pressed too far. Sometimes history of science illuminates and clarifies philosophy of science; but in other cases it only distorts and confuses logical problems if it is forced upon them. For example, the historical fact that J. R. Mayer's contributions to the formulation of the modern energy concept were inspired by observations of blood coloration differences between inhabitants of torrid and temperate zones sheds little light in any philosophic study of the nature of energy. Even more extreme is the case of Kekulé, who discovered the benzene ring in a dream about a serpent biting its tail, an historical vignette clearly irrelevant to the natural philosophy of organic chemistry.

The typical historical account of quantum theory from Planck to the present outlines a rather smooth transition from the "Old Quantum Theory" (Bohr atom, particulate photon, classical ontology) to the "New Quantum Theory" (state vectors, probability, complementarity); and from the purely historical point of view, this evolutionary description is perhaps entirely

acceptable. However, as indicated in the parentheses above, from a philosophical perspective there is no gradual metamorphosis from the "Old" to the "New"; there is an abrupt discontinuity in theoretical structure, hence any discussion about modern quantum theory which employs concepts peculiar to the "Old" to demonstrate alleged features of the "New" is accordingly of little value. And such arguments are not uncommon; in fact, it is a good rule of thumb that any discourse upon quantum theory that calls its content "intuitive" will probably commit this history-inspired blunder of mixing the "Old" with the "New".

Unfortunately, the standard demonstrations of the incompatibility of certain observables in quantum theory are of this type. They are of course the historic gedankenexperiments of Bohr[1] and Heisenberg,[2] which have been both repudiated and defended many times over the past 35 years. The present work will not present still another analysis of these thought experiments. Although such demonstrations are obviously not as irrelevant to the philosophy of quantum theory as is Kekulé's serpent to that of chemistry, nevertheless their primary value is historical, as samples of the motivating thoughts of great physicists engaged in the construction of the quantum theory.

The emphasis in the following sections will be placed rather on a logical study of the notion of compatibility entirely within the axiomatic framework of (New!) quantum theory, independently of whatever dreams, intuitions, or gedankenexperiments historically might have inspired its ingenious creators.

(2) Uncertainty principle: Many gedankenexperiments have been designed to illustrate Heisenberg's famous law; unfortunately, the false impression is often conveyed that his principle, which is actually a theorem about standard deviations in collectives of measurement results, imposes

restrictions on measurability. To see the absurdity of such an inference, consider the following analogous argument. Suppose the widths of a large sample of desks were measured, the results tabulated, and the standard deviation Δw computed; similarly, let Δl, the standard deviation for length measurements, be determined. It would not be especially surprising is there existed a constant k such that $\Delta w \Delta l > k$ even if the sample were extended to include every desk ever manufactured. Now, reasoning as is often done in connection with the uncertainty principle, we would have to conclude from $\Delta w \Delta l > k$ (a) that if the length of a desk is measured to within an "instrumental error" δw, one cannot at the same time contrive a method for measuring the width with less "error" than $\delta l = \frac{k}{\delta w}$, and hence (b) that it is impossible to measure precisely at the same time the length and width of a desk. Obviously, these drastic conclusions are not a logical consequence of the inequality $\Delta w \Delta l > k$. In section 2, the uncertainty principle will be stated correctly in a proper theoretical context and briefly discussed.

(3) Projection postulate (naive version): Frequently appended to the useful postulates of quantum mechanics is one which, if correct, would easily lead to the incompatibility doctrine as a theorem. It is the notion of wave packet reduction, according to which measurement invariably leaves a system in such a state that an immediate repetition of the measurement would yield the same result as the first measurement.[3] It turns out that if simultaneous measurement of noncommuting observables were possible it would usually leave a system in a nonexistent state; thus it is often argued that simultaneous measurement is impossible. This argument is, however, unworthy of serious consideration since the idea of wave packet reduction does not survive close scrutiny. Such reduction cannot be consistently attached to quantum theory by postulation because of the inherent statistical

nature of quantum states; i.e., the physical reference of the density operator to ensembles[4] rationally precludes its changing abruptly in response to a single measurement. Moreover, it is not true in general that an immediately repeated measurement always yields the same result as the first one.[5]

(4) Projection postulate (von Neumann's measurement intervention transformation)[6]: There is a way[7] to express the projection postulate in terms of ensembles and the selection of subensembles which does at least make sense. If this version represented a universal trait of measurement, then it would imply the incompatibility principle as a theorem. We have proved this elsewhere.[8] However, it can be demonstrated that even this "reasonable" variant of the projection postulate does not describe all physical measurements and is therefore unacceptable as a quantal axiom. Hence it is useless as an argument in behalf of incompatibility.

(5) Problems concerning joint probabilities: If joint (i.e., simultaneous) measurements are possible, then there must exist joint probability distributions. However, attempts to generate such distributions for noncommuting observables using fairly standard mathematical ideas have been unsuccessful, and this failure has been interpreted as proof of the incompatibility principle. This position will be examined carefully in section 8.

(6) Von Neumann's simultaneous measurability theorem: In his classic work on quantum mechanics, von Neumann proved a theorem which is undoubtedly the best defense ever given of the incompatibility doctrine. Strangely enough, it is also the most widely ignored argument for incompatibility even though, unlike (1)-(5), it is a logical deduction from a seemingly reasonable quantum axiom set which does not include the projection postulate.* (Cf. sec. 4)

---

*Of course the projection postulate does appear in von Neumann's book, but it plays no role in the theorem here considered.

As stated above, the purpose of the present work is to scrutinize and evaluate the principle of incompatibility of noncommution observables. However, it is not our intention to consider the six preceding arguments one by one; as a matter of fact, most of them will scarcely be mentioned again. In particular, no further discussion will appear concerning (1), (3), and (4). The value of semiclassical gedankenexperiments (1) has already been commented upon; and the absurdity of the naive projection postulate (3) and inadequacy, or lack of universality, of its sometimes correct version (4) are fully discussed elsewhere by the present writer.[4,5] As to the uncertainty principle (2), enough will be said in section 2 to demonstrate its irrelevance to the compatibility principle. Thus ensuing sections will emphasize (5) and (6), the only extant arguments for incompatibility which are firmly embedded in the basic mathematical structure of modern quantum theory.

Because (5) and (6) arise not in shallow classical intuition but deep in the theoretical framework of quantum mechanics, it will be necessary to survey basic quantum axiomatics in order to distinguish clearly which common quantum "truths" are assumed hypotheses and which ones are derivable propositions. Only in this way can the deductions in (5) and (6) be properly evaluated. The remainder of the present section sets the stage for this analysis by reviewing several important definitions from the quantum theory of measurement and then using them to obtain a clear statement of the compatibility problem.

As in other branches of physics, the objects of study in quantum mechanics are called physical systems. Associated with these systems are the constructs known as observables which are in turn correlated via epistemic rules to empirical operations which generate numbers. Such operations are called measurements. The numbers they produce are called measurement

results, and it is the responsibility of quantum theory to regularize, interpret, and make predictions about them. Specifically, quantum physics is designed to cope with problems of this format: given a repeatable laboratory procedure Π for the preparation of physical systems, what will be the statistical distribution of measurement results obtained from measurements performed upon an ensemble of systems all prepared identically (in the manner Π)? This question may refer to any observable and to measurements at any given time after preparation.

To avoid unnecessary philosophical dilemmas, it is extremely important to understand the peculiarly quantum theoretical nexus which relates the concepts of measurement result, observable, and system. Perhaps these connections are best understood by contrasting them to their classical analogues. In prequantum rhetoric the process of measurement could be described as follows: physical systems are endowed with certain observable attributes characterized by numerical values; measurement is an empirical procedure for discovering just what these values are. Thus classically measurement results are simply revelations of the values of observable properties possessed by the system. The key word here is possessed, for it expresses succinctly the classical relationship between measurement results and observables.

In quantum mechanics the connection is a weaker one. It is no longer possible to pictorialize physical systems as objects characterizable by definite values of the observables. A classical billiard ball has values for position, momentum, energy, angular momentum, etc. A "quantum billiard ball" has no such values. Nevertheless, it remains true in quantum physics, as in the rest of science, that a system's observables are operationally defined and that measurements of them do yield the numerical results upon which theory feeds. Thus for the "quantum billiard ball", it is proper to

speak of the numerical results of position, momentum, energy, or angular momentum measurements, but it is improper to interpret these numbers as past, present, or future properties of the ball. According to quantum theory, no physical mode of preparation Π exists which could produce systems certain to yield upon measurement a preassigned pair of values (x,p) for position and momentum, for example.

Incidentally, the possessed quality of classical observables brought the concepts of measurement and preparation conceptually close to one another. Since a measurement operation simply revealed a possessed value, the same operation could also be called a preparation method for obtaining systems having that value of the measured observable. Despite such classical intuition, however, the constructs measurement and preparation must be severed in quantum theory. Failure to do so leads to the projection postulate with its attendant physical and philosophical problems.[9]

Since quantum physics is a theory about measurement results instead of possessed properties, it is natural that a concept of measurement should appear among its primitive terms in a place analogous to that occupied by "possession of attributes" in classical theory. In other words, numbers associated with observables are conceptually linked to systems only through statements like this one: "if observable a is measured on system S, the numerical result a will emerge...". This is the primary meaning of measurement in quantum theory, and we shall designate this measurement construct, which supersedes the classical idea of possession, by the symbol $M_1(a)$.

In quantal as in classical physics, it is also necessary to recognize the basic idea of measurement as an empirical procedure yielding numbers associated with observables. Just as in classical physics one must have both the possession concept and measurement schemes to determine the possessed values, similarly quantum physics requires a traditional measurement

concept in addition to M_1. Assume that in classical and quantum mechanics we have a "direct" operational definition of position. For example, a position measurement may be performed upon a speedometer needle essentially by looking at it. From the position measurement on the needle a number is inferred which is declared to be the result of a speed measurement upon the system of interest. The theoretical justification of such an inference will be called a theory of measurement. A classical theory of measurement for the speedometer would consist of a mathematical analysis of the instrument leading to proof of a correlation between the possessed position values of the needle and possessed speed values of the system. The concept of measurement exemplified by this speed measurement will be denoted by M_2. In short, M_2 is the normal measurement construct of all science; $M_2(\alpha, \beta, \ldots)$ represents any empirical procedure yielding numbers a, b, ... which through a theory can be interpreted as the values associated with observables $\alpha, \beta, \ldots$. In quantum theory, the values are associated with the observables via $M_1(\alpha), M_1(\beta) \ldots$ and not by possession. Thus, for example, a quantum theory of measurement for the speedometer would consist of a mathematical analysis of the instrument leading to proof of a correlation between the probabilities that M_1(needle position) would yield certain values and the probabilities that M_1(speed of system) would yield certain values. A measurement scheme M_2 is thus explained quantum theoretically in terms of the primitive measurement construct M_1.

Using these concepts, it is possible to define precisely what is meant by simultaneous measurability of two observables: observables α and β will be termed compatible, simultaneously measurable, or jointly measurable if there exists an $M_2(\alpha, \beta)$, i.e., an operation yielding two numbers a, b with the same probabilities that quantum theory confers upon the two propositions "$M_1(\alpha)$ yields a" and "$M_1(\beta)$ yields b", where both

$\mathcal{M}_1$'s refer to the same instant in time. The compatibility problem, to which the rest of this paper is devoted, may therefore be stated as follows: if $\mathcal{A}, \mathcal{B}$ are noncommuting observables, is it quantum theoretically possible for an $\mathcal{M}_2(\mathcal{A}, \mathcal{B})$ to exist?

2. Quantum Axiomatics and the Uncertainty Theorem

To provide a framework for systematic analysis of the compatibility problem, it seems appropriate to discuss at the outset certain basic propositions of general quantum axiomatics.* Accordingly, the basic axioms of quantum physics will now be stated, and the proofs of several important theorems will then be reviewed.

P1: (Correspondence Postulate) (Some) linear Hermitean operators on Hilbert space which have complete orthonormal sets of eigenvectors correspond to physical observables. If operator A corresponds to observable $\mathcal{A}$, then the operator $\mathcal{F}(A)$ corresponds to observable $\mathcal{F}(\mathcal{A})$, where $\mathcal{F}$ is a function.

It is convenient to use the symbol $\longleftrightarrow$ to represent this operator-observable correspondence relation; thus $A \longleftrightarrow \mathcal{A}$ means A "corresponds to" $\mathcal{A}$ in the sense of P1. The observable $\mathcal{F}(\mathcal{A})$ is defined by this measurement procedure: measure $\mathcal{A}$ and use the result a to evaluate the given function $\mathcal{F}$; the number $\mathcal{F}(a)$ is then the result of an $\mathcal{F}(\mathcal{A})$-measurement. The function $\mathcal{F}$ of operator A, $\mathcal{F}(A)$, is found by the following standard mathematical procedure: consider the spectral expansion of A, $A = \sum_k a_k P_{\alpha_k}$ where a_k is an eigenvalue and P_{α_k} denotes the projector onto the span of eigenvector α_k; the operator $\mathcal{F}(A)$ is then simply $\sum_k \mathcal{F}(a_k) P_{\alpha_k}$. (Extension to degenerate and/or continuous spectra is straightforward.)

---

*For a fuller discussion of the background concepts, see J. Park, "Quantum Theoretical Concepts of Measurement". (Part II of thesis)

P2: (Mean Value Postulate) To every ensemble of identically prepared systems there corresponds a real linear functional of the Hermitean operators, $m(A)$, such that if $A \leftrightarrow a$, the value of $m(A)$ is the arithmetic mean $\langle a \rangle$ of the results of a-measurements* performed on the member systems of the ensemble.

The content of P1 and P2 is slightly different from that of their analogues in typical von Neumann[10]-inspired axiomatizations. In the original form of the Correspondence Postulate, observables and Hermitean operators were assumed to stand in one-to-one correspondence; in other words, the postulate included both of the following statements: (1) Every observable has an Hermitean operator representative; (2) Every Hermitean operator corresponds to a physical observable. In 1952, Wick, Wightman, and Wigner[11] challenged the symmetry of this quantal correspondence by introducing the concept of superselection rules, i.e., assertions which declare certain Hermitean operators to be unobservable in principle. To embrace superselection rules with minimal theoretic change, the word every in (2) is simply replaced by some: (2') Some Hermitean operators correspond to physical observables. Although this particular proposed variation turns out to be irrelevant in the present context and will henceforth be ignored, it does suggest that the universal correspondence proclaimed in (1) and (2) is less than sacrosanct.

As we shall see later, just as superselection rules challenge the every in (2), an important facet of the compatibility problem hinges on the every in (1). Accordingly, the need will arise subsequently to distinguish between different "degrees" of operator-observable correspondence. For this purpose the following terminology will be adopted: Strong

*In postulates and theorems, the term "a-measurement" refers to the primary quantum measurement construct $M_1(a)$.

correspondence means that both (2') and (1) are assumed; weak correspondence means that the Correspondence Postulate includes (2') but not (1), as in P1 above.

In subsequent sections, the relationship of this choice of correspondence schemes to the problem of compatibility will be developed, and eventually it will be demonstrated that only the weak type (P1) is logically acceptable.

Several"elementary" quantum theorems will now be stated and proved. Although these proofs are not new, we shall nevertheless reproduce them in some detail in order to show their independence from strong correspondence. Strangely enough, although the content of these theorems is well known, the fact that they are theorems, i.e., derivable from P1 and P2, is rarely acknowledged. Instead, they are often given as extra postulates (or conjured up heuristically) in quantum textbooks; such an approach easily gives the impression that P1 and P2 (or the equivalent) are mere guidelines, whereas in fact they rigorously imply all the general propositions of quantum statics.

Th1: For each mean value functional $m(A)$ there exists an Hermitean operator ρ such that for each A,

$$m(A) = Tr(\rho A).$$

Proof:[12] Let $\{\phi_n\}$ be a complete orthonormal set, and $A_{nk} \equiv \langle \phi_n, A\phi_k \rangle$. Since A is Hermitean, $A_{kn} = A^*_{nk}$, i.e., $Re\, A_{kn} = Re\, A_{nk}$, $Im\, A_{kn} = -Im\, A_{nk}$. Since $\{\phi_n\}$ is complete, $\sum_n P_{\phi_n} = 1$. A may therefore be expressed as follows:

$$A = \left(\sum_n P_{\phi_n}\right) A \left(\sum_k P_{\phi_k}\right) = \sum_{kn} P_{\phi_n} A P_{\phi_k} = \sum_{kn} |\phi_n\rangle A_{nk} \langle \phi_k|$$

$$= \sum_n |\phi_n\rangle A_{nn} \langle \phi_n|$$

$$+ \sum_{k<n} (|\phi_n\rangle\langle\phi_k| + |\phi_k\rangle\langle\phi_n|)\, Re\, A_{kn}$$

$$+ \sum_{k<n} i(|\phi_n\rangle\langle\phi_k| - |\phi_k\rangle\langle\phi_n|)\, Im\, A_{kn}.$$

It is easily verified that the following operators are Hermitean:

$$P_{\phi_n} = |\phi_n\rangle\langle\phi_n|,$$
$$K(n,k) = |\phi_n\rangle\langle\phi_k| + |\phi_k\rangle\langle\phi_n|,$$
$$\bar{K}(n,k) = i(|\phi_n\rangle\langle\phi_k| - |\phi_k\rangle\langle\phi_n|).$$

Thus A now has the form of a linear combination of Hermitean operators:

$$A = \sum_n A_{nn} P_{\phi_n} + \sum_{k<n} K(n,k)\,\mathrm{Re}\,A_{kn} + \sum_{k<n} \bar{K}(n,k)\,\mathrm{Im}\,A_{kn}.$$

Now by P2 the mean value functional $m(A)$ is linear. Hence

$$m(A) = \sum_n A_{nn}\, m(P_{\phi_n}) + \sum_{k<n}\left[\mathrm{Re}\,A_{kn}\, m(K(n,k)) + \mathrm{Im}\,A_{kn}\, m(\bar{K}(n,k))\right]$$
$$= \sum_{k,n} \rho_{nk} A_{kn} = \mathrm{Tr}(\rho A),$$

where ρ is defined by its matrix elements:

$$\rho_{nn} \equiv m(P_{\phi_n})$$
$$\rho_{kn} \equiv \tfrac{1}{2} m(K(n,k)) + \tfrac{1}{2} i\, m(\bar{K}(n,k)),\ k<n,$$
$$\rho_{nk} \equiv \tfrac{1}{2} m(K(n,k)) - \tfrac{1}{2} i\, m(\bar{K}(n,k)),\ k<n.$$

Finally, by P2 $m(B)$ is real for any Hermitean B, in particular for $P_{\phi_n}, K(n,k), \bar{K}(n,k)$. Thus $\rho_{nk} = \rho^*_{kn}$; i.e., ρ is Hermitean. Q.E.D.

The operator ρ, known as the statistical operator or density operator, is not only an "index" of measurement statistics, but is also the seat of causality in quantum physics. For this reason, ρ may be called the quantum <u>state</u> of the ensemble to which it refers. The general "law of motion" is given by the following axiom:

P3: (Dynamical Postulate) To every type of closed* quantum system there corresponds a linear unitary operator T (the evolution operator) such that the temporal development of the density operator ρ for an ensemble of such systems is given by

$$\rho(t_2) = T(t_2,t_1)\,\rho(t_1)\,T^\dagger(t_2,t_1).$$

---

*In general, <u>open</u> systems, i.e., those interacting with other systems, do not evolve causally by unitary transformation. Thus composite quantum systems with interacting constituents are in this dynamical sense <u>indivisible</u>. (cf. J. Park, thesis, Part I.)

In the following theorems, we assume the Hermitean operators have discrete spectra; similar propositions hold for the continuous case.

Th2: The probability $W_a(a_k;\rho)$ that an a-measurement on a system from an ensemble with density operator ρ will yield the A-eigenvalue a_k is given by

$$W_a(a_k;\rho) = Tr(\rho P_{\mathfrak{M}_k})$$

where $\mathfrak{M}_k$ is the subspace belonging to a_k.

Proof: $W_a(a_k;\rho)$ is (by the physical definition of probability) the mean value of the observable $\mathcal{F}_k(a)$, where $\mathcal{F}_k$ is defined by

$$\mathcal{F}_k(a) = \begin{cases} 1, & a = a_k \\ 0, & a \neq a_k \end{cases}.$$

By P1, $\mathcal{F}_k(A) \longleftrightarrow \mathcal{F}_k(a)$. Hence, by Th1,

$$W_a(a_k;\rho) = m(\mathcal{F}_k(A)) = Tr(\rho \mathcal{F}_k(A)).$$

Consider the spectral expansion* of A and form $\mathcal{F}_k(A)$:

$$A = \sum_{n d_n} a_n P_{\alpha_{n d_n}};$$

$$\mathcal{F}_k(A) = \sum_{n d_n} \mathcal{F}_k(a_n) P_{\alpha_{n d_n}} = \sum_{d_k} P_{\alpha_{k d_k}} = P_{\mathfrak{M}_k}.$$

$$\therefore W_a(a_k;\rho) = Tr(\rho P_{\mathfrak{M}_k}).$$

Q.E.D.

Th3: $Tr\,\rho = 1.$

Proof: Consider the trivial observable $\mathfrak{1}$ defined as follows: measure any observable a; whatever the result, we shall say that an $\mathfrak{1}$-measurement has been performed with numerical result unity. Symbolically, $\mathfrak{1} = \mathcal{F}(a)$ where $\mathcal{F}(a) = 1$. The operator corresponding to $\mathfrak{1}$ is therefore

$$\mathcal{F}(A) = \sum_{k d_k} \mathcal{F}(a_k) P_{\alpha_{k d_k}} = \sum_{k d_k} P_{\alpha_{k d_k}} = 1,$$

the identity operator.

Now, it is obvious from the definition of $\mathfrak{1}$ that $\langle \mathfrak{1} \rangle = 1$; hence $Tr(\rho 1) = Tr\,\rho = 1.$

Q.E.D.

---

*Here we have allowed for degeneracy; however in later sections spectra will usually be assumed nondegenerate. The eigenvectors belonging to eigenvalue a_n are denoted by $\alpha_{n d_n}$, $d_n = 1, 2, \ldots$

Th4: The only possible results of $\mathcal{A}$-measurements are the eigenvalues $\{a_k\}$ of A, where $A \leftrightarrow \mathcal{A}$.

Proof: The probability Z that an $\mathcal{A}$-measurement will yield a number which is not an A-eigenvalue is equal to the mean value of the observable $\mathcal{Q}(a)$, where

$$\mathcal{Q}(a) = \{ \begin{matrix} 1, \text{ a not an A-eigenvalue} \\ 0, \text{ a an A-eigenvalue} \end{matrix} \}.$$

The operator to which $\mathcal{Q}(a)$ corresponds is

$$\mathcal{Q}(A) = \sum_n \mathcal{Q}(a_n) P_{\psi_n} = \sum_n (0) P_{\psi_n} = 0.$$

Hence $Z = Tr(\rho 0) = 0$. As is customary in scientific applications of probability theory, we assume that zero probability for an event means that the event will never occur*. Q.E.D.

Th5: The density operator ρ is positive semidefinite.

Proof: The projection operator P_ϕ, $P_\phi \leftrightarrow P$, has eigenvalues 0, 1 (ϕ arbitrary). Thus by Th4** $\langle P \rangle \geq 0$. By Th1, $\langle P \rangle = Tr(\rho P_\phi)$. Hence for all ϕ, $Tr(\rho P_\phi) = \langle \phi, \rho \phi \rangle \geq 0$, i.e., ρ is positive semidefinite. Q.E.D.

Although all of the foregoing theorems required only weak correspondence, they would of course still follow if P1 were replaced by an axiom of strong correspondence:

P1S: The set of physical observables is in one-to-one correspondence with the set of linear Hermitean operators on Hilbert space with complete orthonormal sets of eigenvectors. If $A \leftrightarrow \mathcal{A}$, then $\mathcal{F}(A) \leftrightarrow \mathcal{F}(\mathcal{A})$.

---

* From a practical viewpoint, this makes sense; however, in the ideal case of an infinite ensemble, the interpretation of zero probability can at best be that the event in question is of measure zero in the sample space of interest.

**Ignoring superselection, we assume that every P_ϕ represents an observable.

A cursory examination of P1S and P2 seems to suggest that nothing about simultaneous measurement could ever be derived from such axioms, for in them reference is made only to measurements of single observables, i.e., $M_1(a)$. Indeed, the absence of a similar joint measurement construct $M_1(a, B, \ldots)$ appears to justify the conclusion that quantum theory is silent to the problem of compatibility and that in order to discuss simultaneous measurements at all P2 must be augmented by some kind of joint probability postulate. We shall see later (sec. 4) that this "silence" is illusory, that P1S and P2 do in fact place severe restrictions upon simultaneous measurements.

To approach the problem of joint measurements from an axiom set referring only to single measurements, it is necessary to develop a theory of compound observables, i.e., observables defined as functions of several ordinary observables. Then information regarding joint measurements can be extracted from an analysis of single measurements defined as functions of the joint measurement results. For example, a compound observable $\mathcal{F}(a, B)$ may be operationally defined as follows: measure a and B simultaneously, substitute the results a, b into the function $\mathcal{F}(a, b)$; the value $f = \mathcal{F}(a, b)$ is then the result of the $\mathcal{F}(a, B)$-measurement. Then by P1S, there exists an operator F to represent $\mathcal{F}(a, B)$; hence if F is known, $\mathcal{F}(a, B)$-measurements are subject to quantum mechanical analysis, and in this sense joint measurements would be in the domain of the ordinary quantum theory of M_1's.

This leads us directly to an old and interesting quantum problem:[13] given the correspondences $A \leftrightarrow a$, $B \leftrightarrow B$, ... and a compound observable $\mathcal{F}(a, B..)$, what F corresponds to $\mathcal{F}$? Note that if P1S is adopted, the existence of such an F is assured (if $a, B, \ldots$ are simultaneously measurable), for every observable must have an operator representative. If, however, only the weaker P1 holds, the existence of an F such that

$F \longleftrightarrow \mathcal{F}(\alpha, \beta, \ldots)$ is not guaranteed. In neither case is there a general prescription for finding F; but it is obviously necessary to require that all deductions based on a proposed F be consistent with P2, the definition of $\mathcal{F}$, and the theorems reviewed above. In particular, Th1 and Th4 suggest especially useful consistency conditions. To formulate such criteria, the following notation will be helpful.

Let the sets $\mathcal{E}(A)$ and $\mathcal{N}(\mathcal{F})$ be defined as follows: $\mathcal{E}(A)$ is the set of eigenvalues belonging to the operator A; $\mathcal{N}(\mathcal{F})$ is the set of conceivable measurement results associated with an observable $\mathcal{F}$. When $\mathcal{F} = \alpha$, $\mathcal{N}(\alpha) = \mathcal{E}(A)$ by Th4. However, when $\mathcal{F}$ is a function of α and β, for example, it is possible that correlations between α and β might preclude the occurrence of certain a priori conceivable values of $\mathcal{F}$, i.e., preclude certain of the values $\mathcal{F}(a_k, b_\ell)$ calculable from eigenvalues of A and B under the a priori assumption that all eigenvalue pairs (a_k, b_ℓ) are possible. In such a case, $\mathcal{E}(F) \subset \mathcal{N}(\mathcal{F})$. Finally, for a state ρ, let $W(a_k, b_\ell, \ldots; \rho)$ denote the joint probability that simultaneous $\alpha, \beta, \ldots$ -measurements yield $a_k, b_\ell, \ldots$.

Two consistency conditions may now be expressed as follows: If $F \longleftrightarrow \mathcal{F}(\alpha, \beta, \ldots)$, then

(C_1) $\sum_{k\ell} W(a_k, b_\ell, \ldots; \rho)\, \mathcal{F}(a_k, b_\ell, \ldots) = Tr(\rho F)$, for every ρ,

(C_2) $\mathcal{E}(F) \subseteq \mathcal{N}[\mathcal{F}(\alpha, \beta, \ldots)]$.

It is easy to see that (C_1) arises from Th1 and the definition of $\mathcal{F}$ while (C_2) is needed to prevent conflict with Th4. However, the usefulness of (C_1) must be immediately questioned, for the joint probability W is of course unknown. Indeed the search for W is an important phase of the compatibility problem (cf. sec. 8-10). Nevertheless, condition (C_1) is not so mute as it appears, since for the proper choice of $\mathcal{F}$, it becomes independent of the form of W (cf. sec. 4).

It will be noted that both P1 and P1S include explicit postulation of the correspondence $\mathcal{F}(A)\leftrightarrow\mathcal{F}(a)$, and the above survey of the proofs of theorems 1-5 indicates clearly the value of that rule. Nevertheless, later developments could cast doubt upon its necessity as a postulate unless there is good reason to regard it as indispensable. Since Th2 (i.e., the form of W_a) is the cornerstone of practical calculations in quantum theory and is therefore not a proposition which could easily be challenged, the following theorem indicates strongly that the correspondence $\mathcal{F}(A)\leftrightarrow\mathcal{F}(a)$ could not reasonably be removed from the quantum axiom set.

Consistency Theorem: If $W_a(a_k;\rho)=Tr(\rho P_{\ell\ell_k})$ and if there exists an operator F such that $F\leftrightarrow\mathcal{F}(a)$, then $F=\mathcal{F}(a)$, where $A\leftrightarrow a$.

Proof: The operator F must satisfy consistency condition (C_1) :

$$\sum_k Tr(\rho P_{\ell\ell_k})\mathcal{F}(a_k)=Tr(\rho F).$$

Thus $Tr[\rho(F-\sum_k \mathcal{F}(a_k)P_{\ell\ell_k})]=0$ for every ρ, which implies $F=\sum_k \mathcal{F}(a_k)P_{\ell\ell_k}=\mathcal{F}(A)$. (Note that (C_2) is also satisfied.)Q.E.D.

The $\mathcal{F}(a)$ which has probably received more attention than any other is a fairly complicated one: $\mathcal{F}(a)=(a-\langle a\rangle)^2$, where $\langle a\rangle$ is a real constant which is just the arithmetic mean of $M_1(a)$'s on the ensemble of interest. Using P2, we may write $\mathcal{F}(a)=(a-m(A))^2$; then by the correspondence rule in P1, $\mathcal{F}(a)\leftrightarrow(A-m(A)1)^2$. By definition,

$$(\Delta a)^2\equiv m[(A-m(A)1)^2]=\langle(a-\langle a\rangle)^2\rangle;$$

Δa, sometimes called the standard deviation, is a common statistical quantity measured in the obvious way as a function of measurement results from an ensemble.

Historically, Δa has often been linked erroneously to the problem of compatibility by way of what is perhaps the most widely misunderstood theorem in quantum mechanics--the Heisenberg uncertainty principle. This

is not the place to dwell upon its misinterpretations; however, a few remarks are needed in order to dispel the popular contention that the uncertainty principle places restrictions on simultaneous measurability. First of all, precisely what is the uncertainty principle? It is a theorem, rigorously derived from the quantum postulates; it states that under fairly general conditions,

$$\Delta\alpha\,\Delta\beta \geq \tfrac{1}{2}\left|m([A,B])\right|,$$

where A, B are Hermitean operators representing quantum observables α, β and $\Delta\alpha, \Delta\beta$ signify the standard deviations for collectives of α- and β-measurements.

Hence this remarkable theorem has the following physical significance: Given two identical ensembles of quantum systems, if α-measurements are performed on one ensemble, β-measurements on another, then the quantities $\Delta\alpha$, $\Delta\beta$, calculated from the measurement results, will satisfy the relation $\Delta\alpha\,\Delta\beta \geq \tfrac{1}{2}\left|m([A,B])\right|$.

The principal point here stressed is that $\Delta\alpha$ and $\Delta\beta$ have physical meaning only within the context of <u>statistics</u>. It is therefore illogical to interpret the uncertainty principle as a denial of the possibility of simultaneous measurement of α and β upon a single system if $[A,B]\neq 0$, as has sometimes been done. The only sense in which $\Delta\alpha\,\Delta\beta$ may refer to a single system is purely <u>statistical</u>, i.e., to an ensemble involving one system sequentially measured and reprepared. Furthermore it should be noted that $\Delta\alpha\,\Delta\beta$, as presented above, is not even calculated from <u>simultaneous</u> measurements of α and β performed on each system. Whether or not that is a meaningful alternative method of calculation remains at this point undecided and will be deferred to later sections. For the present, it suffices to observe that whatever conclusions are reached concerning the notion of compatibility, i.e., simultaneous measurability of several

observables on a single system, there can be no conflict with the uncertainty principle, a relation involving statistical properties of measurements of single observables.

Finally the uncertainty principle is often interpreted not as a denial of simultaneous measurability but as a statement about the accuracy of simultaneous measurements of non-commuting observables. For the interesting special case of position X and momentum P (whose operators satisfy $[X,P]=i\hbar 1$) this interpretation typically runs as follows: It is impossible to measure simultaneously X and P exactly; the product of the "inaccuracies" $\Delta X \Delta P$ is never less than what the uncertainty principle allows, i.e., $\Delta X \Delta P \geq \frac{1}{2}\hbar$. Properly understood as physicists' jargon, this interpretation of the principle is not too objectionable, for its implicit meaning is the same as the more careful explanation above. To see this, it is only necessary to realize that the phrase "to measure simultaneously X and P exactly" here refers experimentally to a collective of X- and P- measurement results, each obtained by a measurement performed upon a member of the ensemble at some given time relative to the preparation of that member. From such a collective the term "exact" draws its physical meaning; thus if X-measurements are performed on an ensemble of identically prepared systems each at the same time relative to preparation, and if all these results are identical, experimental jargon would say that "an exact position measurement has been made," since $\Delta X = 0$.

To summarize: whatever propositions about joint measurements may or may not be consistently incorporated into quantum theory, the uncertainty principle remains unscathed so long as its interpretation does not transcend the content of its theoretical statement and proof by making unjustified references to joint, rather than single, measurements. Conversely, the uncertainty principle is not an a priori restriction on any consideration

purely about joint measurements; as noted earlier, perhaps the principle can be generalized to cover joint measurements once a theory for the latter has been devised, but no such requirement need be stipulated in advance. Indeed, the uncertainty principle is irrelevant to the problem of compatibility.

3. Trivial Joint Measurements and Commutativity

There is one type of joint measurement whose consistency with quantum theory is certain, for it involves the performance of only one measurement upon the system. The resulting number is then used to generate a set of numbers through a set of established functions; hence, the simultaneous measurement of a set of observables has been performed, albeit in a rather trivial sense. Accordingly, joint measurements performed simply by arithmetical manipulation of one measurement result for a single observable will henceforth be called trivial joint measurements.

The question then arises as to whether the joint measurement of any two observables is reducible to a trivial joint measurement; if so, quantum theory could embrace the concept of simultaneous measurement in a very natural way. However, the correspondence rule $\mathcal{F}(A) \leftrightarrow \mathcal{F}(a)$ may be used to prove that any two operators jointly measurable in this trivial sense necessarily commute. To see this, assume the existence of an observable c and functions $\mathcal{F}, \mathcal{G}$ such that the observables a, B are expressible in the form

$$a = \mathcal{F}(c), \quad B = \mathcal{G}(c).$$

By the above correspondence rule, if $C \leftrightarrow c$, $C\gamma_k = c_k \gamma_k$,

$$\mathcal{F}(c) \leftrightarrow \mathcal{F}(C) = \sum_k \mathcal{F}(c_k) P_{\gamma_k},$$

$$\mathcal{G}(c) \leftrightarrow \mathcal{G}(C) = \sum_n \mathcal{G}(c_n) P_{\gamma_n}.$$

Since two different operators cannot correspond to the same observable*, it follows that $A = \mathcal{F}(C)$, $B = \mathcal{G}(C)$. Hence

$$[A,B] = \sum_{kn} \mathcal{F}(c_k)\mathcal{G}(c_n)[P_{ok}, P_{on}] = 0.$$

The trivial joint measurements thus do not exhaust all a priori conceivable simultaneous measurements. But this does not imply that noncommuting observables are incompatible; it merely establishes that they are not trivially compatible. Nevertheless, since $[A,B]=0$ is (1) a necessary condition for trivial joint measurability of $\mathcal{A}$ and $\mathcal{B}$ and (2) the only condition under which $\Delta\mathcal{A}\Delta\mathcal{B}=0$ may hold, it is sometimes claimed via a misinterpretation of the uncertainty principle that the only simultaneous measurements permitted by quantum theory are the trivial ones, that commutativity is the mathematical criterion of compatibility. However, in view of our preceding remarks about the uncertainty principle, such a position is evidently illogical.

Although the notion of trivial joint measurement is not an adequate basis for a general treatment of simultaneous measurements, it does provide a means for deriving the joint probabilities associated with several commuting observables. (The problem of joint distributions for noncommuting observables will be explored in sections 8-10)

If $\mathcal{A}$ and $\mathcal{B}$ are simultaneously measurable through an auxiliary observable $\mathcal{C}$, then joint probabilities for the results of $\mathcal{M}_1(\mathcal{A})$ and $\mathcal{M}_1(\mathcal{B})$ are calculable through single probabilities associated with $\mathcal{M}_1(\mathcal{C})$. Given A, B satisfying the necessary condition $[A,B]=0$, the problem is to find $\mathcal{F}, \mathcal{G}$ and $\mathcal{C} \leftrightarrow C$ such that

$$\mathcal{A} \leftrightarrow A = \mathcal{F}(C) \leftrightarrow \mathcal{F}(\mathcal{C}),$$
$$\mathcal{B} \leftrightarrow B = \mathcal{G}(C) \leftrightarrow \mathcal{G}(\mathcal{C}).$$

---

*If $\mathcal{A} \leftrightarrow A_1$ and also $\mathcal{A} \leftrightarrow A_2$, then for every ρ, $\langle a \rangle = Tr(\rho A_1) = Tr(\rho A_2)$; but this implies $A_1 = A_2$.

By a mathematical theorem generally omitted from standard quantum textbooks, the condition $[A,B]=0$ is sufficient to insure the existence of $\mathcal{F}, \mathcal{G}$, and C such that

$$A=\mathcal{F}(C),\ B=\mathcal{G}(C).$$

While some cases of the general theorem[14] are rather complicated, it is not difficult to prove the theorem for operators A, B which have discrete spectra. Consider, for example, the interesting case where A and B constitute what Dirac called a complete set of commuting observables. This means that the set of common eigenvectors (the existence of which $[A,B]=0$ assures) can be labeled so that the eigenvalue equations take this form:

$$A\,\gamma_{nm}=a_n\,\gamma_{nm},$$
$$B\,\gamma_{nm}=b_m\,\gamma_{nm}.$$

Now, let $\{c_{kl}\}$ denote a set of distinct real numbers and define $\mathcal{F}, \mathcal{G}$ by the relations

$$a_n=\mathcal{F}(c_{nl}),\ b_m=\mathcal{G}(c_{km})$$

This determines the desired C; its spectral expansion is just

$$C=\sum_{kl} c_{kl}\,P_{\gamma_{kl}}.$$

Hence $$\mathcal{F}(C)=\sum_{kl}\mathcal{F}(c_{kl})P_{\gamma_{kl}}=\sum_{kl}a_{k}\,P_{\gamma_{kl}}=A,$$
$$\mathcal{G}(C)=\sum_{kl}\mathcal{G}(c_{kl})P_{\gamma_{kl}}=\sum_{kl}b_{l}\,P_{\gamma_{kl}}=B.$$

(Note that $\mathcal{F}, \mathcal{G}$, and C are not unique.)

To find the joint probability $W(a_k, b_l; \rho)$ that $M_1(a)$ and $M_1(B)$ will yield (a_k, b_l) for the state ρ, we simply calculate $W_c(c_{kl};\rho)$ the probability that $M_1(c)$ will yield c_{kl}, by the standard quantum mechanical rule for single probabilities:

$$W(a_k, b_l;\rho)=W_c(c_{kl};\rho)=\mathrm{Tr}(\rho\,P_{\gamma_{kl}}).$$

The non-uniqueness of $\mathcal{F}, \mathcal{G}$, and C is now seen to be unproblematical; $W(a_k, b_l;\rho)$ does not depend on them. (As always happens in quantum theory, in the continuous spectrum generalization W becomes w, a probability

density.) It is easy to illustrate this theorem by deriving the familiar joint co-ordinate distribution of nonrelativistic wave mechanics. Consider the function space whose functions (vectors) have the form $\phi(x,y)$. The operators corresponding to co-ordinates $\mathcal{X}$ and $\mathcal{Y}$ are postulated to be multiplication by x and y, respectively; but it is not necessary to postulate further that $\phi^*\phi$ is the joint probability density for $M_1(\mathcal{X})$ and $M_1(\mathcal{Y})$. Since $[X,Y]=0$, we may conclude that $\mathcal{X}$ and $\mathcal{Y}$ are jointly measurable in the trivial sense through an auxiliary observable $\mathcal{C} \leftrightarrow C$; the eigenvectors of C will be $\gamma_{x_1,y_1}(x,y) = \delta(x-x_1)\delta(y-y_1)$. From the generalization of the preceding theorem to include continuous spectra, it follows that the joint probability density for $\mathcal{X}$ and $\mathcal{Y}$ in the (pure) state $\rho = P_\phi$ is given by

$$\begin{aligned} w(x_1,y_1;P_\phi) &= w_{\mathcal{C}}(c_{x_1,y_1};P_\phi) = |\langle \gamma_{x_1,y_1}, \phi\rangle|^2 \\ &= \left|\int dx \int dy\, \delta(x-x_1)\delta(y-y_1)\phi(x,y)\right|^2 \\ &= \phi^*(x_1,y_1)\phi(x_1,y_1). \end{aligned}$$

For this simple example it is possible to see empirically the difference between trivial and nontrivial joint measurements. Consider a plane fluorescent screen upon which a rectangular cartesian co-ordinate system has been established. Whenever a glowing spot appears on the screen, a joint measurement of $\mathcal{X}$ and $\mathcal{Y}$ for the impinging particle may be made by reading the $\mathcal{X}$ and $\mathcal{Y}$ scales separately. Two numbers thus emerge, the $\mathcal{X}$- and $\mathcal{Y}$-measurements having been performed nontrivially. To construct an apparatus which simultaneously measures $\mathcal{X}$ and $\mathcal{Y}$ in the trivial sense, we require an observable $\mathcal{C}$ and the functions $\mathcal{F}, \mathcal{G}$. A suitable observable $\mathcal{C}$ may be defined operationally by assigning to every point on the fluorescent screen a single real number. When a particle strikes the screen, the single number co-ordinated with the glowing spot is regarded as the result of a $\mathcal{C}$-measurement upon the particle.

If the functions $\mathcal{F}$ and $\mathcal{G}$ define a mapping from the single number to the ordinary (x,y)-co-ordinates of the point*, the entire scheme then constitutes a trivial joint measurement of $\mathcal{X}$ and $\mathcal{Y}$.

The following basic assumption about the simultaneous measurement of commuting observables underlies the joint probability theory just outlined: (J_1) If it is possible to measure $\mathcal{A}$ and $\mathcal{B}$ simultaneously in the trivial sense, i.e., if $[A,B]=0$, then the joint probability $W(a_k, b_\ell; \rho)$ calculated using an auxiliary observable $\mathcal{C}$ is valid for all simultaneous $\mathcal{A}, \mathcal{B}$ -measurements regardless of whether or not they are actually performed by means of $\mathcal{C}$ -measurements. Strangely enough, this assumption apparently never receives explicit statement though many normal applications of quantum mechanics would be difficult to justify without it.

As we have seen above, even the derivation of the standard interpretation of wave functions depends upon (J_1). Indeed all correlations among quantum observables are ultimately based on that assumption.

It is instructive to restate the content of (J_1) in a mathematically definite manner which seems less ad hoc: (J_2) The joint probability $W(a_k, b_\ell; \rho)$, $[A,B]=0$, is a unique functional of the state ρ. Expressed in this way, the assumption is quite reasonable; it merely requires that the state of an ensemble be sufficient to determine the distribution, as would be the case in classical physics. In particular, no additional information regarding the method of measurement is needed to obtain W; thus once a W for a given ρ is found by the method of trivial joint measurement, it is naturally assumed that this W is the W associated with the given ρ independently of how $\mathcal{A}$ and $\mathcal{B}$ might be measured.

---

*It is true that such a mapping defies ordinary geometric intuition; nevertheless, it does exist, for the line and plane are of the same order of infinity, viz., that of the continuum.

Suppose, however, that $[A,B]\neq 0$. Then the method of trivial joint measurements is of course inapplicable. Does (J_2) still hold? Is the quantum state ρ alone sufficient to determine W's for joint measurements of noncommuting observables? We shall study this matter in sections 8-10.

Finally, it is perhaps of some interest to know that the postulate (J_1) or (J_2) has the following trivial consequence: two simultaneous measurements of the same observable a must each yield the same result. Formally, we may choose $c=a, \mathcal{F}(c)=\mathcal{G}(c)=a$. Assignment of probabilities to $\mathcal{F}(c)$- and $\mathcal{G}(c)$-measurements via ordinary quantal analysis of c-measurements obviously leads to nonzero probability only when "both" a-measurements yield identical results, viz., that of the auxiliary $c(=a)$-measurement.

4. Von Neumann's Theorem: Noncommuting Observables Are Incompatible

The popular belief that the only compatible observables are the trivially compatible ones was reviewed in section 2, where the uncertainty principle, the standard basis of this dogma, was presented and found irrelevant. However, there exists also a rather formidable logical demonstration that if two observables are compatible they are trivially compatible. It is an elegant theorem[15] due to von Neumann which, strangely enough, appears to be almost universally ignored, even by proponents of the viewpoint for which it is the strongest support. Indeed the main impact of the theorem seems to have been to influence mathematicians[16] interested in modern physics to define the term "simultaneously measurable" by the commutativity condition for trivial joint measurability, which is not very helpful in view of the fact that both words in common physical usage already had other definitions, as explained in section 1. Because von Neumann's theorem is of central importance to the problem of compatibility, it is appropriate here

to scrutinize it carefully, paying special attention to the hypotheses on which it is based.

Let A and B denote two Hermitean operators corresponding to observables $\mathcal{A}$ and $\mathcal{B}$. Assume that $\mathcal{A}$ and $\mathcal{B}$ are compatible (jointly measurable). If $\mathcal{A}$ and $\mathcal{B}$ are simultaneously measured on a system, two numbers will result. Now, suppose those numbers are added (subtracted) and the sum (difference) is considered to be the result of measuring an observable $\mathcal{S}(\mathcal{D})$ expressed symbolically by $\mathcal{S}=\mathcal{A}+\mathcal{B}$ $(\mathcal{D}=\mathcal{A}-\mathcal{B})$. Clearly if $\mathcal{A}$ and $\mathcal{B}$ are compatible observables, $\mathcal{S}$ and $\mathcal{D}$ are observables. Therefore, in accordance with a widely accepted quantal postulate (P1S) there must exist Hermitean operators S and D representing observables $\mathcal{S}$ and $\mathcal{D}$. It is easy to prove that if S exists, S = A + B.

To do this, recall the consistency criterion (C_1) of section 2 which the operator S must satisfy:

$$(C_1)\quad \sum_{kl} W(a_k, b_l; \rho)\, \mathcal{F}(a_k, b_l) = Tr(\rho S), \text{ for every } \rho.$$

Here $\mathcal{F}(a, B) = a + B$; therefore

$$\sum_k \Big[\sum_l W(a_k, b_l; \rho)\Big] a_k + \sum_l \Big[\sum_k W(a_k, b_l; \rho)\Big] b_l = Tr(\rho S).$$

Because of the additive form of $\mathcal{F}$, we see that the unknown joint probability $W(a_k, b_l; \rho)$ may now be replaced by the well known quantally prescribed marginal probabilities:

$$\sum_k W(a_k, b_l; \rho) = W_B(b_l; \rho) = Tr(\rho P_{\beta_l}),$$

$$\sum_l W(a_k, b_l; \rho) = W_a(a_k; \rho) = Tr(\rho P_{\alpha_k}).$$

Hence, for every ρ,

$$\sum_k Tr(\rho P_{\alpha_k}) a_k + \sum_l Tr(\rho P_{\beta_l}) b_l = Tr(\rho S).$$

This determines S uniquely, as follows:

$$\begin{aligned} Tr(\rho S) &= Tr\Big(\rho \sum_k a_k P_{\alpha_k}\Big) + Tr\Big(\rho \sum_l b_l P_{\beta_l}\Big) \\ &= Tr(\rho A) + Tr(\rho B) \\ &= Tr[\rho(A+B)], \text{ for every } \rho. \end{aligned}$$

Therefore $S = A + B$.

Similarly, $D = A - B$.

From the correspondence rule $\mathcal{F}(a) \leftrightarrow \mathcal{F}(A)$ of section 2 together with those just derived there follows another:

$$\mathcal{A}\mathcal{B} \leftrightarrow \tfrac{1}{2}(AB + BA)$$

where $\mathcal{A}\mathcal{B}$ denotes the observable measured by multiplying the results of simultaneous measurements of $\mathcal{A}$ and $\mathcal{B}$. The argument proceeds as follows. If $\mathcal{A}$ and $\mathcal{B}$ are compatible, $\mathcal{S}$ and $\mathcal{D}$ are also, as well as $\mathcal{S}^2$ and $\mathcal{D}^2$. Thus $\mathcal{R} = \frac{1}{4}(\mathcal{S}^2 - \mathcal{D}^2)$ is an observable, and by the just mentioned correspondences, its operator must be $R = \frac{1}{4}(S^2 - D^2)$. To see the meaning of this strange operator, suppose that a joint measurement of $\mathcal{A}$ and $\mathcal{B}$ yields the numbers a and b. Proper manipulation of a and b yields a number r, which is by definition the result r of an $\mathcal{R}$-measurement: $r = \frac{1}{4}[(a+b)^2 - (a-b)^2] = ab$. Hence, the operator for the product $\mathcal{A}\mathcal{B}$ must be

$$R = \tfrac{1}{4}[(A+B)^2 - (A-B)^2] = \tfrac{1}{2}(AB + BA).$$

Von Neumann's simultaneous measurability theorem is based on the correspondence rule just derived. It should be noted that neither this rule nor the ones from which it follows are arbitrary postulates; they are all deduced from the axioms, as demonstrated just above and in section 2. We shall now state von Neumann's theorem and review the proof since it is not well known.

Simultaneous Measurability Theorem: If $\mathcal{A}$ and $\mathcal{B}$ are compatible and $\mathcal{A} \leftrightarrow A$, $\mathcal{B} \leftrightarrow B$, then

$$[A, B] = 0.$$

Proof: Since $\mathcal{A}$ and $\mathcal{B}$ are compatible, it follows that any function of $\mathcal{A}$ and $\mathcal{B}$ is observable (simply by measuring $\mathcal{A}$ and $\mathcal{B}$ simultaneously and using the results to evaluate the function); in particular, consider the function $\mathcal{A}^2\mathcal{B}$. By the correspondence rules discussed

above, $a^2 \leftrightarrow A^2$ and hence

$$a^2 B \leftrightarrow \frac{A^2 B + B A^2}{2} \equiv P_1$$

But the product correspondence rule also implies

$$a^2 B = a(aB) \leftrightarrow \frac{1}{2}\left(A \frac{AB+BA}{2} + \frac{AB+BA}{2} A\right)$$
$$= \frac{1}{4}(A^2 B + 2ABA + BA^2) \equiv P_2 .$$

Since quantum theory cannot tolerate the ambiguity of having several operators for a single observable (cf. fn., p.22), we must set $P_1 = P_2$ and accept the consequences. The resultant equation is

$$ABA = \frac{1}{2}(A^2 B + BA^2).$$

Moreover, the same argument is applicable to any functions $\mathcal{F}(a)$ and $\mathcal{G}(B)$; thus

(1) $\mathcal{F}(A)\mathcal{G}(B)\mathcal{F}(A) = \frac{1}{2}[\mathcal{F}^2(A)\mathcal{G}(B) + \mathcal{G}(B)\mathcal{F}^2(A)]$.

For simplicity consider A and B having discrete spectra (the general case is not essentially different): $A\alpha_k = a_k \alpha_k$, $B\beta_l = b_l \beta_l$.

Define functions $\mathcal{F}_k$ and $\mathcal{G}_l$ as follows:

$$\mathcal{F}_k(a) = \delta_{a a_k}, \quad \mathcal{G}_l(b) = \delta_{b b_l}.$$

Then
$$\mathcal{F}_k(a) \leftrightarrow \mathcal{F}_k(A) = \sum_n \delta_{a_n a_k} P_{\alpha_n} = P_{\alpha_k},$$
$$\mathcal{G}_l(B) \leftrightarrow \mathcal{G}_l(B) = \sum_m \delta_{b_m b_l} P_{\beta_m} = P_{\beta_l}.$$

Substituting these operators into condition (1), we have for each k, l,

$$P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \frac{1}{2}\left(P_{\alpha_k}^2 P_{\beta_l} + P_{\beta_l} P_{\alpha_k}^2\right).$$

But $P_{\alpha_k}^2 = P_{\alpha_k}$, $P_{\beta_l}^2 = P_{\beta_l}$.

Hence (2) $P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \frac{1}{2}(P_{\alpha_k} P_{\beta_l} + P_{\beta_l} P_{\alpha_k})$.

Multiply (2) on the right by P_{α_k}:

(3) $P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \frac{1}{2}(P_{\alpha_k} P_{\beta_l} P_{\alpha_k} + P_{\alpha_k} P_{\beta_l} P_{\alpha_k})$.

Multiply (2) on the left by

(4) $P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \frac{1}{2}(P_{\alpha_k} P_{\beta_l} P_{\alpha_k} + P_{\beta_l} P_{\alpha_k})$.

Subtract (4) from (3):

$$P_{\alpha_k} P_{\beta_l} - P_{\beta_l} P_{\alpha_k} = 0.$$

Therefore $[A,B]=\sum_{\kappa\rho} a_\kappa b_\rho [P_{\alpha\kappa}, P_{\beta\rho}]=0$

This completes the proof of von Neumann's simultaneous measurability theorem.

Expressed succintly, it says that if $\mathcal{A}$ and $\mathcal{B}$ are compatible, they are trivially compatible, for their operators necessarily commute. Unlike the semi-classical gedankenexperiments, the vague interpretations of the uncertainty principle, and some strange philosophizing about subjective wave packet reductions, the foregoing theorem offers an argument strong and clear in behalf of the proposition that non-commuting observables cannot even in principle be measured simultaneously. It affirms that the very notion of general compatibility simply cannot logically be appended to the established theoretical structure of quantum physics, unless the latter is somehow modified. This possibility of nullifying the theorem by such a basic alteration in the quantum postulates will be considered later.

Evidently inspired by the preceding theorem, various authors[17] have suggested that quantum mechanics should be rephrased in a new logical framework which would properly allow for incompatibility. We intend to show in subsequent sections that von Neumann's mathematics in fact does not establish incompatibility as an intrinsic quantal property but rather proves something else. Hence, if our analysis is correct, any "quantum logic" designed to embrace incompatibility is motivated by and founded upon a mistaken interpretation of quantum physics. Accordingly, we shall not review such a system in any detail; however, it is instructive to expose certain salient features of "quantum logic" to establish clearly its relation to von Neumann's theorem.

Propositions, or questions, are easily introduced into quantum theory as functions of observables. For example, consider an observable $\mathcal{A}\leftrightarrow A=\sum_\kappa a_\kappa P_{\alpha\kappa}$ and the proposition $\mathcal{P}_n$: "$\mathcal{M}_1(\mathcal{A})$ would yield a_n."

The proposition P_n is just the observable measured as follows: measure a; if a_n results, assign P_n the value 1 (by convention); if $a_{k\neq n}$ results, assign P_n the value 0. In short, $P_n = f_n(a)$ where f_n is defined by $f_n(a_k) = \delta_{a_n a_k}$. Hence $P_n \longleftrightarrow f_n(A) = \sum_{k=1}^{N} f(a_k) P_{a_k} = P_{a_n}$.

Similarly, a suitable projection operator may be found for any proposition involving commuting observables; but because of von Neumann's theorem, any compound proposition involving noncommuting observables must of course be regarded as undecidable, or absurd. For any two compatible propositions, P and Q, it is possible to find operators corresponding to the logical relations P "or" $Q \equiv P \cup Q$ and P "and" $Q \equiv P \cap Q$:

$$P \cup Q \longleftrightarrow P + Q - PQ,$$
$$P \cap Q \longleftrightarrow PQ.$$

The change in logic said to be necessitated by quantum mechanics has to do with the classical distributive law of propositions:

$$P \cap (Q \cup R) = (P \cap Q) \cup (P \cap R).$$

Suppose A, B are operators in a two dimensional Hilbert space. If $[A,B] \neq 0$, and $P \longleftrightarrow P_{\beta_1}$, $Q \longleftrightarrow P_{a_1}$, $R \longleftrightarrow P_{a_2}$, then because of von Neumann's theorem, the distributive law cannot hold in quantum theory. To see this, note that

$$P \cap (Q \cup R) \longleftrightarrow P_{\beta_1}(P_{a_1} + P_{a_2} - P_{a_1} P_{a_2}) = P_{\beta_1}(1-0) = P_{\beta_1};$$

but $(P \cap Q) \cup (P \cap R)$ is an absurd proposition because neither $P \cap Q$ nor $P \cap R$ is measurable since they are compounds of A and B with $[A,B] \neq 0$. Thus, since the distributive law apparently cannot hold in quantum theory, it has been suggested that "nondistributive" logic is required for modern physics.

Contrary to this view, we shall later show that once von Neumann's theorem is properly interpreted, quantum theory provides no reason to eradicate the distributive law of logic.

5. Counterexamples: Noncommuting Observables Are Not Incompatible

Mathematically speaking, von Neumann's simultaneous measurability theorem is beyond criticism. Neither logical nor algebraic errors are involved; the theorem is a legitimate deduction from P1S and P2. If, therefore, one could find a counterexample, i.e., describe quantum mechanically a physical process fully certifiable as a simultaneous measurement of, say, position and momentum, then the basis of von Neumann's theorem would require reformulation. It would then establish not the incompatibility of physical observables but rather the inconsistency of the quantum mechanical axioms. In this section we shall discuss the construction of such counterexamples.

It is instructive to consider first a typical quantum theory of measurement for a single observable, viz., the "time-of-flight" method for measuring the momentum P of an electron. We assume that the rule of correspondence for position X is an ideal one which might consist, for example, of the direct observation of a coincidence between a scale marking and a macroscopic spot appearing on a photographic plate in response to an electron impact.

Suppose we have an "electron gun" which prepares the state $\rho = P_\psi$. Using nonrelativistic wave mechanics, the probability density $W_\rho(p;\psi)$ for $M_1(P)$ at the time of preparation is easily calculated:

$$W_\rho(p;\psi) = |\langle \pi_p, \psi\rangle|^2 = \left|\int_{-\infty}^{\infty} dx \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{ipx}{\hbar}} \psi(x)\right|^2,$$

where $P \leftrightarrow P = \frac{\hbar}{i}\frac{\partial}{\partial x}$, $\pi_p = (2\pi\hbar)^{-\frac{1}{2}} \exp \frac{ipx}{\hbar}$, $P\pi_p = p\pi_p$, $\langle \pi_{p_1}, \pi_{p_2}\rangle = \delta(p_1 - p_2)$. This distribution is the quantum mechanical touchstone for deciding whether or not a proposed experiment which generates numbers via the established operational definition for $M_1(X)$ qualifies as a momentum measurement scheme $M_2(P)$. If the numbers in question are to be regarded as $M_1(P)$ results, they must satisfy the theoretical distribution $W_\rho(p;\psi)$.

Let $t = 0$ be the time when the electron is known to be in the prepared state $\rho = P_\psi$. The wave function $\psi(x, t=0)$ is assumed to be of compact support, and it is convenient to set up the origin of the x-axis so that the interval where $\psi(x) \neq 0$ is $(-x_0, x_0)$. The $M_2(P)$-procedure[18] is simple: wait a very long time $(t \to \infty)$ as the electron moves <u>freely</u>, then measure the observable $F(x) = \frac{mx}{t}$, where m is the electron mass. The resultant number then counts as the result of $M_1(P)$ at $t = 0$. To justify this operational definition of P quantum mechanically, we must prove that the probability for $M_1(P)$ to yield $\rho\varepsilon(p_1, p_2)$ at $t = 0$ equals the probability that $M_1[F(x)]$ yields $\frac{mx}{t}\varepsilon(p_1, p_2)$ as $t \to \infty$. In short, it must be shown that

$$W_\rho[\rho\varepsilon(p_1, p_2); \psi(x, 0)] = W_{F(x)}\left[\frac{mx}{t}\varepsilon(p_1, p_2); \psi(x, t), t \to \infty\right].$$

To find $\psi(x, t)$, given $\psi(x, 0)$, one must use the general quantum theoretical "law of motion". In the present case, $\rho = P_\psi$ and the law of motion reduces to $\psi_t = T(t, 0)\psi_0$, where T is the free evolution operator for the electron. In function space, this transformation is given by

$$\psi(x, t) = \sqrt{\frac{m}{2\pi i\hbar t}}\exp\frac{imx^2}{2\hbar t}\int_{-\infty}^{\infty}\exp\left[-\frac{im}{\hbar t}\left(xx_1 - \frac{x_1^2}{2}\right)\right]\psi(x_1, 0)\,dx_1.$$

Now, $W_{F(x)}\left[\frac{mx}{t}\varepsilon(p_1, p_2); \psi(x, t)\right] = W_x\left[x\varepsilon\left(\frac{p_1 t}{m}, \frac{p_2 t}{m}\right); \psi(x, t)\right]$

$$= \int_{\frac{p_1 t}{m}}^{\frac{p_2 t}{m}} |\psi(x, t)|^2 dx = \frac{m}{2\pi\hbar t}\int_{\frac{p_1 t}{m}}^{\frac{p_2 t}{m}} dx \left|\int_{-\infty}^{\infty}\exp\left[-\frac{im}{\hbar t}\left(xx_1 - \frac{x_1^2}{2}\right)\right]\psi(x_1)\,dx_1\right|^2.$$

Changing the variable of integration by letting $x = \frac{pt}{m}$, we obtain

$$W_{F(x)}\left[\frac{mx}{t}\varepsilon(p_1, p_2); \psi(x, t)\right] = \int_{p_1}^{p_2}\frac{dp}{2\pi\hbar}\left|\int_{-\infty}^{\infty}\exp\left[\frac{im}{2\hbar t}x_1^2 - \frac{ipx_1}{\hbar}\right]\psi(x_1)\,dx_1\right|^2.$$

Since $\psi(x_1) = 0$ for $x_1 \notin (-x_0, x_0)$, the integral over x_1 is just equal to $\int_{-x_0}^{x_0}\exp\left[\frac{im}{2\hbar t}x_1^2 - \frac{ipx_1}{\hbar}\right]dx_1$, and we may take the limit $t \to \infty$ inside the integral without difficulty. The result is then

$$W_{F(x)}\left[\frac{mx}{t}\varepsilon(p_1, p_2); \psi(x, t), t \to \infty\right] = \int_{p_1}^{p_2}\left|\int_{-\infty}^{\infty}\frac{e^{-\frac{ipx_1}{\hbar}}}{\sqrt{2\pi\hbar}}\psi(x_1)\,dx_1\right|^2 dp.$$

But this equals $\int_{p_1}^{p_2} w_Q[p;\psi(x,0)]\,dp = W_Q[p\in(p_1,p_2);\psi(x,0)]$.

Hence the results of "direct" $\mathcal{F}(x)$-measurements performed sufficiently long after the preparation of $\psi(x,0)$ will be distributed just like the theoretical results for $M_1(P)$ upon $\psi(x,0)$. This time-of-flight arrangement is thus fully certified quantum mechanically as an operational definition of P (for pure states $\psi(x,0)$ of compact support). Because quantum theory can make only statistical predictions, no further guarantee that this method "really" makes P-measurements is required. Indeed further <u>quantal</u> analysis of the question is theoretically inconceivable. Of course it is possible to note that a free classical particle initially confined to $(-x_0,x_0)$ would, by Liouville's theorem, after time t have momentum in the range

$$p = \frac{mx}{t} \pm \frac{mx_0}{t}$$

and that as $t\to\infty$, $p\to\frac{mx}{t}$. However, strictly speaking, this simple classical demonstration adds nothing to the quantal argument just given. A serious theory of measurement should not rely on heuristic classical analogies; it should establish its consistency wholly on quantum mechanical grounds by matching probabilities in the manner illustrated.

With this understanding, it is easy to see that this time-of-flight method for obtaining the result that $M_1(P)$ at $t = 0$ would yield also determines the result that $M_1(P)$ at any time $t > 0$ would yield. This follows from the fact that momentum is conserved in the free motion of the electron; in quantum mechanical terms,

$$W_Q[p\in(p_1,p_2);\psi(x,0)] = W_Q[p\in(p_1,p_2);\psi(x,t)].$$

Thus by the same reasoning which validated the time-of-flight method as a rule of correspondence for $M_1(P)$ at $t=0$, we can likewise regard the results of $M_1[\mathcal{F}(x)]$, $t\to\infty$, as $M_1(P)$-results for <u>any</u> $t > 0$. In particular, consider that instant when the electron strikes the photographic

plate and the result emerges. For that instant we may conclude with full quantum mechanical justification that $M_1(P)$ would have yielded $\psi(x)$ where x is the result of the $M_1(x)$. Hence, contrary to the prohibitions of von Neumann's theorem, we have an empirical method for the simultaneous measurement of x and P, two noncommuting observables!

Incidentally, it should not be thought that the physical unattainability of $t \to \infty$ vindicates the incompatibility doctrine by preventing the performance of a "perfect" time-of-flight P-measurement. Although it is true even classically that time-of-flight P-measurements are never perfect for finite t (unless x at $t = 0$ is known exactly), nevertheless classically and quantally the error inherent in the method can be reduced below any arbitrary limit simply by choosing sufficiently large t, as shown above. Besides, the $t \to \infty$ approximation is not a special property of (x,P)-measurements in general; for example, a magnetic deflection method for joint measurement of x and P which does not require $t \to \infty$ will be outlined later (sec. 10).

There is a tendency among interpreters of quantum theory to dismiss simultaneous measurement schemes such as the one just described as if they did not in fact legitimately challenge the orthodox view that x and P for example, cannot be measured simultaneously. The usual argument seems to have been originated by Heisenberg and may be summarized by his statement[19] that "the uncertainty relation does not refer to the past". In the time-of-flight experiment, for example, by the time the x,P-values emerge, the time to which they refer--the instant just prior to the electron's collision with the photographic plate--is past; and the electron is then buried in the plate. According to Heisenberg[20] such "knowledge of the past is of a purely speculative character, since it can never...be used as an initial condition in any calculation of the future progress of the

electron and thus cannot be subjected to experimental verification. It is a matter of personal belief whether such a calculation concerning the past history of the electron can be ascribed any physical reality or not".

In rejoinder to this distinctly philosophical argument, we offer the following comments:

(1) Knowledge is a dangerous word to employ in discussions regarding quantum measurement. From a strict quantal point of view an electron never possesses properties q, p, etc., of which one can conceivably be knowledgeable or ignorant. There simply does not exist a preparation scheme Π which produces electrons always yielding the same q, p-values from q, p-measurements; the relation $\Delta q \Delta p \geq \frac{\hbar}{2}$ among standard deviations is simply a quantitative expression of this fundamental fact. Accordingly, measurement should never be described as though it increased knowledge by revealing the actual, previously unknown, "value" of an observable. Measurements simply generate numerical results associated with certain operations and observations upon the system of interest. The meaning of these numbers is provided by the theory into which they are fed; in quantum theory it happens that the numbers are not to be regarded as measures of possessed attributes.

(2) It is therefore pointless to say that the uncertainty relations do not refer to the past. They refer to the standard deviations of collectives of measurement results at any time. What the relations do not refer to is measurements upon a single system at a single time; standard deviations naturally refer only to measurements upon ensembles. Hence, as explained in section 2, the emergence of simultaneous q, p-values upon measurement in no way violates the uncertainty principle.

(3) In the time-of-flight method, the q, p-measurement results refer of course to the instant just prior to the electron's impact in the plate.

These numbers are indeed useless for predicting in classical style the result of a future X-measurement, but they are no more "speculative" or lacking in "physical reality" than any other measurement result. Their lack of predictive power stems from the fact that the "motion" of quantum systems is not governed by Newtonian laws. That the X, P-values refer to a past time is no special property of <u>simultaneous</u> measurements; it is characteristic of all quantum measurements. The time-of-flight measurement of P alone referred to $t = 0$ although the result did not emerge until $t \to \infty$. Nevertheless, such P-measurements play a key role in the process of empirical verification; for example, their statistical distribution determines whether or not the state prepared by the "electron gun" is really $\psi(x,0)$. Indeed, if the physical significance of X, P-values is a matter of "personal belief", then all measurement results for <u>single</u> observables are likewise of solipsistic value only.

We therefore conclude that the foregoing method for simultaneous measurement of X, P is as significant as any other quantum mechanical measurement scheme. And its validity cannot be philosophized away.

The time-of-flight method for $M_2(P)$ exhibits a curious feature which seems at first paradoxical. Only X is "directly" measured; P is then measured by calculation of $\mathcal{F}(X)$. Yet if $P = \mathcal{F}(X)$, it follows (cf. section 2) for the respective operators that $P = \mathcal{F}(X)$, hence $[X, P] = 0$, which is false. To see this more accurately, it is convenient to describe the time-of-flight method in the Heisenberg picture, where $X_0 \leftrightarrow X$, $P_0 \leftrightarrow P$, $X_t \leftrightarrow X_t$, $P_t \leftrightarrow P_t$, $X_t = T^\dagger X T$, $P_t = T^\dagger P T$, T being the <u>free</u> evolution operator. In these terms, it would appear that the time-of-flight scheme is based on the relation $P = \frac{m}{t} X_t$, $t \to \infty$. But from this expression and momentum conservation $(P = P_t)$, it follows that

$$[X, P] = [T^{+}XT, T^{+}PT] = [T^{+}XT, P] = [X_t, \frac{mX_t}{t}] = 0,$$

which contradicts $[X, P] = i\hbar 1$.

The solution to this dilemma provides interesting information regarding quantum measurement. The error in the above reasoning inheres in the assumption that $P = \frac{mX_t}{t}$ just because $P = \frac{mX_t}{t}$ is used in the time-of-flight arrangement. Actually, the relation $P = \frac{mX_t}{t}$ holds only for initial wave functions $\psi(x,0)$ of compact support that develop freely until the measurement of X_t, $t \to \infty$. In general, $[X, P] \neq 0$ and no $\mathcal{F}$ exists such that $P = \mathcal{F}(X)$. Nevertheless, this analysis reveals two interesting points concerning quantum measurements: (1) Even if $[A, B] \neq 0$ so that no <u>general</u> trivial joint measurement scheme for A and B can be constructed, it may still be possible, for certain states and measurement arrangements, to measure B as a function of A (or vice versa). (2) Conversely, the fact that $B = \mathcal{F}(A)$ for some <u>particular</u> $M_2(B)$ does not imply $B = \mathcal{F}(A)$. However, as shown in section 2, if $B = \mathcal{F}(A)$ always, then the operator relation $B = \mathcal{F}(A)$ is valid.

To conclude this section, we introduce another counterexample to the simultaneous measurability theorem. Consider two quantum systems S_1, S_2 with observables A_1, B_1 and A_2 associated with S_1 and S_2, respectively. Suppose $[A_1, B_1] \neq 0$ and denote eigenvectors and eigenvalues as follows:

$$A_1 \alpha_R^{(1)} = a_R^{(1)} \alpha_R^{(1)}, \quad A_2 \alpha_R^{(2)} = a_R^{(2)} \alpha_R^{(2)}.$$

Let $S_1 + S_2$ be noninteracting but in a correlated state:

$$\Psi = \sum_R c_R \alpha_R^{(1)} \otimes \alpha_R^{(2)}$$

If A_2 has an established operational definition, the correlation in Ψ which relates $M_1(A_1)$-results to $M_1(A_2)$-results may be exploited in the standard way to construct an $M_2(A_1)$. As in the time-of-flight $M_2(P)$ case, we must establish a theoretical matching between probabilities associated with $M_1(A_2)$ and $M_1(A_1)$. Since $[A_1, A_2] = [A \otimes 1, 1 \otimes A_2] = 0$,

A_1 and A_2 may be jointly measured (trivially) through an auxiliary observable (cf. section 3). The joint probability $W(a_k^{(1)}, a_l^{(2)}; \Psi)$ is therefore easily calculated:

$$W(a_k^{(1)}, a_l^{(2)}; \Psi) = Tr(P_\Psi P_{a_k^{(1)}} \otimes a_l^{(2)}) = |C_{kl}|^2 \delta_{kl}.$$

From this expression it is apparent that when $M_1(A_2)$ yields $a_k^{(2)}$ a simultaneous $M_1(A_1)$ would yield $a_k^{(1)}$. Hence we have an $M_2(A_1)$ scheme: to measure A_1, simply measure A_2; if $a_k^{(2)}$ results, then $a_k^{(1)}$ is regarded as the result of $M_1(A_1)$.

Suppose B_1, like A_2, has an established operational definition. Now, since the $M_2(A_1)$ just outlined involves no interaction with S_1, we may perform $M_1(B_1)$ simultaneously with $M_2(A_1)$, and thereby jointly measure noncommuting observables A_1 and B_1. Once again von Neumann's theorem is contradicted.

6. Strong Correspondence--the Axiomatic Root of Quantal Inconsistencies

The necessary deductions to be made from the last two sections may be summarized as follows: (1) The quantal postulates (PlS, etc.) rigorously imply that noncommuting observables are incompatible. (2) The same postulates together with what would seem to be a normal scientific understanding of the term measurement may be used to describe empirical arrangements which must be regarded as legitimate schemes for the simultaneous measurement of noncommuting observables. (3) Hence the standard postulates of quantum theory are inconsistent. We must therefore re-examine the axiomatic basis of von Neumann's simultaneous measurability theorem and isolate, if possible, the false hypothesis which enabled the rigorous deduction of this false theorem.

As explained earlier (sections 2, 4) in order to derive a theorem about simultaneous measurement of several observables from axioms referring

only to single observable measurements $M_1(a)$, the notion of compound observable had to be introduced, subject to certain consistency conditions, (C_1) and (C_2) which would have to be satisfied by any operator corresponding to such a compound observable. Condition (C_1) was in general useless because it involved unknown joint probability distributions. However, in the special case of the sum of two observables $a+B$, the latter condition assumed a simple form and, moreover, it even sufficed to <u>determine</u> uniquely the correspondence $a+B \leftrightarrow A+B$ upon which the simultaneous measurability theorem was ultimately based. In fact, once this correspondence is established, the logic of the theorem cannot be doubted, as careful restudy of its proof (section 4) will demonstrate.

We therefore direct attention to the correspondence $a+B \leftrightarrow A+B$. As just noted, condition (C_1) alone <u>implied</u> this rule. To be more explicit, PlS guaranteed the existence of an operator corresponding to the observable $a+B$; that operator would necessarily satisfy (C_1) and (C_2). It then turned out that (C_1) for $a+B$ could be satisfied by only <u>one</u> operator, A + B. <u>Thus condition (C_2) was never used.</u> This observation provides an important clue in our search for the false hypothesis which made possible the proof of von Neumann's (false) simultaneous measurability theorem.

Is $a+B$ really an observable? If not, PlS cannot be invoked to assure the existence of an operator counterpart. To establish the observability of $a+B$, we need only recall the last counterexample of section 5, which showed how two noninteracting but correlated systems S_1 and S_2 could be used to construct an appropriate rule of correspondence for simultaneous $M_1(a_1)$ and $M_1(B_1)$. Since the experimenter is obviously free to add the two results, it is apparent that a_1+B_1 is indeed observable. <u>Therefore, by PlS, there must exist an operator S such that</u> $a_1+B_1 \leftrightarrow S$.

For simplicity, let system S_1 be a "spin" whose relevant states and operators refer to a two-dimensional spinor space. For noncommuting observables A_1, B_1, take x- and z-components of spin $\mathcal{S}_x$ and $\mathcal{S}_z$. In terms of the Pauli operators, we have

$$A_1 = \mathcal{S}_x \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B_1 = \mathcal{S}_z \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Now, simultaneous measurements of A_1 and B_1 using the correlation scheme involving auxiliary system S_2 will by Th4 always yield one of these eigenvalue pairs: $(\frac{\hbar}{2}, \frac{\hbar}{2}), (\frac{\hbar}{2}, -\frac{\hbar}{2}), (-\frac{\hbar}{2}, \frac{\hbar}{2}), (-\frac{\hbar}{2}, -\frac{\hbar}{2})$. (The two eigenvalues of any Pauli matrix are 1, -1.) Hence upon addition to obtain $A_1 + B_1$-measurement results, only the three values, $\hbar, 0, -\hbar$ are possible. To use the set notation of section 2,

$$\mathcal{N}(A_1 + B_1) = \{-\hbar, 0, \hbar\}$$

and by consistency condition (C_2), if $A_1 + B_1 \leftrightarrow S$, quantum mechanics would be self-contradictory unless

$$(C_2) \quad \mathcal{E}(S) \subseteq \mathcal{N}(A_1 + B_1)$$

But (C_1) must also be satisfied by S and, as shown in section 4, the <u>only</u> S meeting this requirement is, for the $A_1 + B_1$ of the present example,

$$S = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Now, by an elementary calculation, the eigenvalues of this operator are $\frac{\hbar}{\sqrt{2}}, \frac{-\hbar}{\sqrt{2}}$; i.e., $\mathcal{E}(S) = \{-\frac{\hbar}{\sqrt{2}}, \frac{\hbar}{\sqrt{2}}\}$. Comparing the sets $\mathcal{N}(A_1+B_1)$ and $\mathcal{E}(S)$, $A_1 + B_1 \leftrightarrow S$, we find that

$$\mathcal{N}(A_1 + B_1) \cap \mathcal{E}(S) = \emptyset.$$

Thus the only operator S capable of satisfying (C_1) does not satisfy (C_2).

To summarize: $A_1 + B_1$ is demonstrably observable. P1S then insures the existence of $S \leftrightarrow A_1 + B_1$. If the quantal axioms are consistent, that S must satisfy both (C_1) and (C_2). The <u>unique</u> S which satisfies (C_1) violates (C_2). Hence, the quantal axioms P1S and P2 are inconsistent.

This is of course the conclusion reached at the beginning of this

section upon confrontation of von Neumann's theorem with the counterexamples of section 5. This time, however, we have isolated the difficulty within the initial hypotheses of that theorem and are now able to understand "why" the theorem is false. The reason is simply that P1S--strong correspondence--proclaims the existence of operator-observable correspondences which simply cannot exist in harmony with the remaining postulates. Thus the axiom set--P1S, P2--must be altered.

In view of the overwhelming empirical success of quantum mechanics, it is immediately clear that any proposed axiomatic modification should, if possible, be a "slight" one; i.e., it should remove the inconsistencies attached to the problem of compatibility, but it should not affect the normal applications of the quantal algorithm by revising or eliminating any common theoretical procedures. As we have seen, the troublesome correspondences which breed inconsistencies are derived essentially from two quantal propositions: (1) P1S, invoked to guarantee the existence of an operator to represent any given observable, and (2) Th1, which establishes the general form of the quantum mean value functional, viz., $Tr(\rho A)$. Obviously, we cannot change just Th1 without also altering the postulates; moreover, since that theorem is the basis for the highly successful quantum theoretical state representations (the density operators), it is, practically speaking, not a reasonable candidate for deletion or even revision. We are left therefore only with P1S to criticize.

In section 2, a distinction was drawn between the usual axiom of strong correspondence (P1S), which claims a one-to-one relation between observables and Hermitean operators, and a simpler axiom P1, called weak correspondence, which states only that every Hermitean operator represents an observable. It should be clear that P1 omits just that part of P1S which led to the inconsistencies discussed above. This suggests that P1,

not P1S, should be adopted as the true quantal Correspondence Postulate. The question as to whether this truncation of P1S to P1 still permits the derivation of key quantal propositions such as Th1 will be deferred to the next section, where our advocacy of weak correspondence will be fully defended.

If P1S is replaced by P1, what does von Neumann's theorem really prove? It becomes a reductio ad absurdum proof that the correspondence $\alpha+\beta \leftrightarrow A+B$ and its consequence $\alpha\beta \leftrightarrow \frac{1}{2}(AB+BA)$ are not valid if $[A,B]\neq 0$. To be specific, in the proof of the theorem, at the stage where two distinct operators emerge to represent one observable, one faces two logical possibilities: (1) Interpret this ambiguity as the failure of strong correspondence, or (2) equate the two operators to derive a condition on A and B for the existence of the compound observables involving α and β, in short, a condition for their simultaneous measurability. As presented in detail earlier, von Neumann chose the second alternative, which enabled his rigorous derivation of $[A,B]=0$ as the condition of simultaneous measurability.

But what of the other possibility? Is it not perhaps more reasonable to take alternative (1) and to regard the conclusion from (2) as an absurdity which shows that no operators correspond to $\alpha+\beta$ or $\alpha\beta$ when $[A,B]\neq 0$? Of interest in this connection is an old but apparently forgotten, mathematical objection to the principle of strong correspondence, a theorem due to Temple.[21]

The basic premise is strong correspondence. As we have seen, it follows that if $A\leftrightarrow\alpha, B\leftrightarrow\beta$ then

$$A+B\leftrightarrow \alpha+\beta\,,\ \tfrac{1}{2}(AB+BA)\leftrightarrow\alpha\beta.$$

Now, when these rules together with $\mathcal{F}(A)\leftrightarrow\mathcal{F}(\alpha)$ are applied to the observable $\alpha\beta\varepsilon$, there results a threefold ambiguity:

$$(aB)c \longleftrightarrow \tfrac{1}{2}\left[\tfrac{1}{2}(AB+BA)C + \tfrac{1}{2}C(AB+BA)\right]$$
$$= \tfrac{1}{4}(AB+BA)C + \tfrac{1}{4}C(AB+BA);$$
$$(ca)B \longleftrightarrow \tfrac{1}{4}(CA+AC)B + \tfrac{1}{4}B(CA+AC);$$
$$(Bc)a \longleftrightarrow \tfrac{1}{4}(BC+CB)A + \tfrac{1}{4}A(BC+CB).$$

Once again, we face two alternatives: (1) Interpret the ambiguity as the failure of strong correspondence, or (2) use the ambiguity to derive a relation among A, B, C. Temporarily choosing (2), it may be readily shown that equating the above operators for aBc leads to the relation $[C,[A,B]]=0$, for every A, B, C.

Up to this point, Temple's theorem seems to be nothing but a variation on von Neumann's theorem. There is a radical difference, however, in the final conclusion to be drawn. From the condition $[C,[A,B]]=0$ and its permutations, it follows in several steps* that A, B, C constitute a commuting set; hence, since a, B, c denoted <u>any</u> physical observables, any two operators which represent physical observables must commute! Needless to say, such a statement violently clashes with the most successful parts of quantum theory, and cannot therefore be permitted to stand. In short, it must be regarded as absurd, and the hypotheses from which it is derived immediately fall into doubt. Yet the similarity of reasoning in Temple's and von Neumann's theorems is striking. Indeed the von Neumann conclusion--the incompatibility of non-commuting observables--should likewise be considered <u>absurd</u>; and the common root of both of these perplexing

---

*$[C,[A,B]]=0$ implies $[A,B]=d(A,B)1$, $d(A,B)$ a number, for any operators A, B. Then, $\frac{1}{2}A[B,C]=\frac{1}{2}d(B,C)A$ and $\frac{1}{2}[B,C]A=\frac{1}{2}d(B,C)A$; adding these equations gives $\frac{1}{2}(ABC-ACB+BCA-CBA)=d(B,C)A$. Similarly, $-\frac{1}{2}(BCA-BAC+CAB-ACB)=d(A,C)B$. Adding the last two equations yields

$$d(B,C)A + d(A,C)B = d\left(\tfrac{AB+BA}{2},C\right)1, \quad \text{for every A, B, C.}$$

Hence

$$d(B,C)=d(A,C)=d\left(\tfrac{AB+BA}{2},C\right)=0, \quad \text{for every A, B, C.}$$

theorems, viz., strong correspondence, should be abandoned. The benefits and liabilities of such a structural change in quantal foundations will be examined in the following section.

7. The Consequences of Weak Correspondence

The suggestion that strong correspondence be abandoned is not altogether welcome, primarily because quantum theory would suffer a certain loss of universality. No longer would every physical observable be automatically considered a quantal observable; there would be conceptual room for nonquantal observables, to which some might even attach the controversial adjective "hidden". Moreover, the pillar of faith upon which the search for rigorous mappings from classical to quantal observables is based would be gone; but acceptance of consequences such as these is not really an insurmountable task. On reflection, it is just as easy to imagine that strong correspondence is probably false as to presume for the sake of completeness that it is probably true. For example, one might speculate half-seriously that the power of the infinity (in Cantor's sense) of physical observables probably exceeds that of all Hermitean operators.

More significant for the physicist is the nonspeculative question: what effect does the replacement of strong correspondence by weak correspondence as a quantum axiom have on the principal quantum theorems? Consider, for example, Th1, which states (in part) that every real linear functional $m(A)$ on the Hermitean operators may be expressed in the form $Tr(\rho A)$. Such a mathematical theorem is quite independent of the physical problem as to whether operators can be found to represent all observables; all that matters is that the operators which <u>are</u> involved do represent observables. Within the mathematical framework--which is of course the context wherein theorems are proved--operational definitions are

irrelevant, and quantum mechanics is just a set of mathematical objects subject to given rules. (Among these are tacit rules concerning the construct $\mathcal{M}_1(a)$ which give meaning to the primitive term observable as it appears in P1.) Thus none of the developments in linear algebra which form the theoretical structure of quantum theory will be affected by the elimination of strong correspondence. In fact a careful search through quantum theory by the present writer for a proposition dependent upon strong correspondence revealed that no basic theorem involving the analysis of ensembles, statistics of measurement results, etc. requires P1S rather than P1 in its proof. (E.g., cf. Theorems 1-5 in sec. 2.)

As mathematical intuition suggests, the only type of theorem which would require strong correspondence is of this general form: "the operator corresponding to the observable $\mathcal{F}$ measured by...is F." Here strong correspondence is effectively standing in the background and commanding the operator algebra to produce an F satisfying that set of requirements which comprises the definition of $\mathcal{F}$, but leaving no way to save the integrity of the theory in the event that no such F exists.

On the other hand, if only weak correspondence is adopted, no theorem of the foregoing kind is ever contemplated. Given an observable $\mathcal{F}$, the operator algebra is not expected to produce an F; instead, it is simply asked whether or not F does exist such that $\mathcal{F} \leftrightarrow F$. In short, what were formerly regarded as "correspondence theorems" are now interpreted as tests of validity for proposed correspondences. Here is a summary of the correct interpretation of the theorems of this kind which were presented in previous sections:

(1) $a+B \leftrightarrow A+B$: (C_1) uniquely determines the operator A + B but (C_2) is often violated. The correspondence is therefore not generally valid.

(2) $\mathcal{A}\mathcal{B} \leftrightarrow \frac{1}{2}(AB+BA)$: Von Neumann's "simultaneous measurability" theorem is merely a demonstration that this correspondence can apply only to commuting operators (in which case it takes the simple form $\mathcal{A}\mathcal{B} \leftrightarrow AB=BA$.)

(3) $\mathcal{A}\mathcal{B}\mathcal{C} \leftrightarrow ?$: Temple's theorem is further proof of the inconsistency of the correspondence (2).

Incidentally, for $[A,B] \neq 0$, rules (1) and (2) are just special cases of the postulated correspondence $\mathcal{F}(\mathcal{C}) \leftrightarrow \mathcal{F}(C)$. To see this, recall that $[A,B]=0$ implies the existence of $C, \mathcal{F}, \mathcal{G}$ such that $A=\mathcal{F}(C)$, $B=\mathcal{G}(C)$. Hence $\mathcal{A}+\mathcal{B} = \mathcal{F}(\mathcal{C})+\mathcal{G}(\mathcal{C}) \leftrightarrow \mathcal{F}(C)+\mathcal{G}(C)=A+B$ and $\mathcal{A}\mathcal{B}=\mathcal{F}(\mathcal{C})\mathcal{G}(\mathcal{C}) \leftrightarrow \mathcal{F}(C)\mathcal{G}(C)=AB=BA$. In this connection it is instructive to consider a simple example which illustrates why consistency condition (C_2) required only $\mathcal{E}(F) \subseteq \mathcal{N}[\mathcal{F}(\mathcal{A},\mathcal{B})]$ rather than $\mathcal{E}(F)=\mathcal{N}[\mathcal{F}(\mathcal{A},\mathcal{B})]$. Let $\mathcal{A}=\mathcal{L}_z^2$, $\mathcal{B}=\hbar\mathcal{L}_z$, where $\mathcal{L}_z$ is the z-component of orbital angular momentum, $\mathcal{L}_z \leftrightarrow L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}$. Suppose $\mathcal{A}$ and $\mathcal{B}$ are measured simultaneously and the results added together. $\mathcal{N}(\mathcal{A}+\mathcal{B})$, the set of all a priori conceivable results of this procedure is given by $\mathcal{N}(\mathcal{L}_z^2+\hbar\mathcal{L}_z)=\{m^2\hbar^2+n\hbar^2\}$, since $\mathcal{E}(L_z)=\{m\hbar\}$. Now, the eigenvalues of $L_z^2+\hbar L_z$ comprise the set $\mathcal{E}(L_z^2+\hbar L_z)=\{k(k+1)\hbar^2\}$ which is only a subset of $\mathcal{N}(\mathcal{A}+\mathcal{B})$, i.e.,

$$\mathcal{E}(L_z^2+\hbar L_z) \subset \mathcal{N}(\mathcal{L}_z^2+\hbar\mathcal{L}_z).$$

The reason this set inequality appears is easily understood if postulate (J_1) or (J_2) of section 3 is recalled. Any measurement of the observables $\mathcal{L}_z^2$ and $\hbar\mathcal{L}_z$ must yield results correlated in the same manner as would be the results of a <u>trivial</u> joint measurement of these observables. One such trivial joint measurement would involve simply measuring $\mathcal{L}_z$ and evaluating $\mathcal{L}_z^2+\hbar\mathcal{L}_z$. Obviously, this procedure could yield only numbers in the set $\{k(k+1)\hbar^2\}=\mathcal{E}(L_z^2+\hbar L_z)$. This demonstration

merely affirms the consistency of (J_1) or (J_2) with the postulated correspondence $\mathcal{F}(\mathcal{C}) \leftrightarrow \mathcal{F}(C)$.

Unfortunately, elementary treatments of quantum mechanics occasionally employ correspondences (1) and (2) as if they represented a universal method of "deriving" quantum operators from classical functions. Since (1) and (2) are, for most $\mathcal{A}, \mathcal{B}$, false, it is evident that so-called "quantization" schemes based upon (1) and (2) are in fact nothing but memory aids which presuppose familiarity with classical mechanics.

For example, consider a <u>classical</u> harmonic oscillator, i.e., a system whose energy is given by the function

$$\mathcal{W}(x,p) = \frac{p^2}{2m} + \frac{1}{2}kx^2, \; k>0.$$

Supposedly, the oscillator is "quantized" by using the established correspondences $\mathcal{X} \leftrightarrow X, \mathcal{P} \leftrightarrow P$ plus rules (1) and (2) to find operator H such that $\mathcal{W} \leftrightarrow H$. This procedure yields immediately

$$H = \frac{1}{2m}P^2 + \frac{k}{2}X^2,$$

the energy operator which then serves to define the <u>quantum</u> harmonic oscillator. To see that this scheme is merely mnemotechnical and has no significant logical value, it is sufficient to note that the classical <u>function</u> $\mathcal{W}$ contributed nothing to the <u>operator</u> H except its functional form. If the definition of the quantum energy operator were based in a <u>physical</u> sense upon the classical energy function, then the operational definition of energy which $\mathcal{W}(x,p)$ entails would also be valid for the quantum energy operator: quantum energy would like classical energy be measurable by first measuring observables $\mathcal{X}, \mathcal{P}$, then using the results x, p to evaluate $\mathcal{W}(x,p)$. Thus the quantum energy operator H would be defined such that $\mathcal{W}(\mathcal{X}, \mathcal{P}) \leftrightarrow H$ i.e., H would have to satisfy consistency conditions (C_1) and (C_2).

Now it turns out that $H = \frac{1}{2m}P^2 + \frac{k}{2}X^2$

does not meet these requirements. The logical situation is therefore the following: the quantum system called "simple harmonic oscillator" is by definition one whose energy observable $\mathcal{H}$ corresponds to the operator above, the form of which can be remembered by classical analogy. $\mathcal{H}$ has its own rules of correspondence, but among these the classical prescription--measure $\mathcal{Q}$ and $\mathcal{P}$, evaluate $W(x,p)$--does not appear. It is quite possible to measure $\mathcal{Q}$ and $\mathcal{P}$ simultaneously and obtain thereby a value for $W(x,p)$ which is not an eigenvalue of H. For example, assume the oscillator is system S_1 in the joint measurement scheme outlined at the end of section 5. Let the correlation between S_1 and S_2 (which are noninteracting) be such that a momentum measurement upon S_2 determines the potential result of a concurrent $M_1(P)$ on S_1. Consider the S_1-subensemble defined by the property that such $M_1(P)$ would have yielded $p=p_1$; this subensemble is therefore characterized by the eigenfunction $\psi_{p_1} = (2\pi\hbar)^{-1} e^{\frac{ip_1 x}{\hbar}}$ (δ-function normalized), which obviously assigns equal relative probability to all results of $M_1(Q)$. Hence this kind of simultaneous Q,P-measurements upon the oscillator will often lead to values of $W(x,p)$ which are not in the set $\mathcal{E}(H) = \{(n+\frac{1}{2})\hbar\sqrt{\frac{k}{m}}\}$. Hence by Th4 the classical energy function $W(x,p)$ has in a sense nothing to do with the quantum energy observable $\mathcal{H} \leftrightarrow H$.

Let $w(x,p)$ be the joint probability density associated with the simultaneous measurements of $\mathcal{Q}$ and $\mathcal{P}$. Our analysis of the relation between $W(x,p)$ and H indicates clearly that we should not expect $w(x,p)$ to satisfy the following equation*:

$$\int dx \int dp\, w(x,p) W(x,p) = Tr(\rho H).$$

*Another example to illustrate this point may be constructed by contrasting measurements of the observable $\mathcal{H}+\frac{1}{2}\mathcal{X}^2$ to those of the observable whose operator is $H+\frac{1}{2}X^2$. The former can lead to any result, the latter only to numbers in the set $\{(n+\frac{1}{2})\hbar\sqrt{\frac{k}{m}}\}$. Thus the operator $H+\frac{1}{2}X^2$ and the observable $\mathcal{H}+\frac{1}{2}\mathcal{X}^2$ do not refer to the same physical situation.

This observation is of special importance in connection with the problem of joint probability in quantum theory (section 8).

With the replacement of P1S by P1, the theoretical structure of quantum mechanics is freed of the various inconsistencies exposed in earlier sections. Fortunately, this simple axiomatic change has no effect whatsoever on the normal applications of the theory to experiment, such as the calculation of scattering cross sections, spectral intensities, etc. But this revision does have considerable theoretical and philosophical significance. Indeed, the elimination of these logical inconsistencies from quantum theory automatically removes the only sound theoretical foundation the concept of incompatibility ever had. Von Neumann's simultaneous measurability theorem is now recognized as a correct mathematical theorem physically misinterpreted as a restriction on measurability; as we have seen, it is in fact a reductio ad absurdum proof that the correspondence $\alpha\beta \leftrightarrow \frac{1}{2}(AB+BA)$ is false unless $[A,B]=0$, or in other words, a proof that $[A,B]=0$ is a necessary condition for the validity of $\alpha\beta \leftrightarrow \frac{1}{2}(AB+BA)$.

Hence any physical or metaphysical idea motivated by or founded upon the concept of incompatibility now requires careful re-examination. Three common incompatibility-based propositions are the following: (1) Because noncommuting observables are in principle not simultaneously measurable, it is meaningless to contemplate joint probability distributions of quantal measurement results; (2) Since any proposition about the results of simultaneous measurements of noncommuting observables is meaningless, a new system of logic is required for quantum physics; and (3) The operators corresponding to two local observables separated by a space-like interval must commute ("microcausality"). These inferences from the false hypothesis of incompatibility will now be discussed in sequence.

(1) The incompatibility doctrine having been discarded, there remains no a priori restraint against the study of joint distributions. Accordingly, in the remaining sections of this work, quantum joint probabilities will be studied systematically.

(2) At the end of section 4, we indicated how incompatibility led to the notion that quantum mechanics requires a new, "nondistributive" logic, i.e., a system which does not involve the law,

$$P \cap (Q \cup R) = (P \cap Q) \cup (P \cap R),$$

which merely expresses an idea most physicists--including quantum theorists--regard as "common sense." The problem was that propositions $P, Q,$ and R can be given for which there does exist an Hermitean operator corresponding to the left member but there is not one for the right member. Apart from the esoteric context in which it is cast, this problem is not different from the difficulty encountered with the correspondence $A+B \leftrightarrow S$. Just as an appropriate S exists only when $[A,B]=0$ similarly a D exists such that $P \cap Q \leftrightarrow D$ only when $[P,Q]=0$. When $[P,Q] \neq 0$, it simply means that the compound proposition $P \cap Q$ has no operator representative D. Naturally it is then impossible to write down an operator counterpart to the distributive law; but this does not make the law wrong! Thus when PlS is replaced by Pl, it becomes apparent that the search for "quantum logics" receives no legitimate motivation or endorsement from quantum physics.

(3) The "microcausality" principle[22] of relativistic quantum field theory is often introduced as though it somehow combined in one grand statement the principles of relativity and complementarity. Actually the rather naive argument involved is nothing but a pun on the term "interference". Consider two local observables $A(x_1), B(x_2)$, i.e., observables associated with space-time points x_1, x_2. According to the theory of

relativity, no signal can connect χ_1 and χ_2 if they are separated by a space-like interval in Minkowski space. Hence no operation at χ_1 could possibly "interfere" with an operation at χ_2. On the other hand, if $[A(\chi_1), B(\chi_2)] \neq 0$, then the orthodox version of quantum theory insists that $A(\chi_1)$ and $B(\chi_2)$ are not simultaneously measurable, or as this is often colorfully phrased, any attempt to measure noncommuting observables simultaneously will be obstructed by mutual "interference" of the measurement procedures employed. Therefore, if χ_1 and χ_2 are space-like separated, relativity then precludes any "interference" between the measurement operations; hence $[A(\chi_1), B(\chi_2)]=0$ necessarily. Clearly this argument is grounded in the canonical misinterpretation of commutativity as an index of measurability. The argument is therefore untenable; it should be understood, however, that no pretense is here made of disproving the "microcausality" principle itself. Its generalization to quantum field operators plays a major role in proofs of the TCP and spin-statistics theorems, and the mathematical property of commutators which the term "microcausality" represents may indeed be a necessary physical postulate. What we have established is that "microcausality" should be regarded as a new postulate, devoid of elementary quantum physical motivation and rather undeserving of its suggestive name.

8. Joint Probability in Quantum Theory

Research concerning joint probability distributions of noncommuting quantum observables has been reported by various authors.[23] Interestingly, their motivations seem to rest upon rather diverse problems --hidden variables, operator-observable correspondence, reduction of diagonal matrix elements to phase space integrals. However, the common feature of all attempts to derive joint distributions has been failure. In general, the

procedure is to set up equations which any proper joint distribution must satisfy and then attempt to solve these equations for the distribution. The natural starting point for such an investigation is the consistency condition (C_1) : if $\mathcal{F}(\mathcal{A},\mathcal{B}) \leftrightarrow F$ and $W(a_k, b_\ell; \rho)$ is the (unknown) joint probability that $M_1(\mathcal{A}), M_1(\mathcal{B})$ would yield a_k, b_ℓ, respectively, then

$$\sum_{k\ell} W(a_k, b_\ell; \rho)\, \mathcal{F}(a_k, b_\ell) = Tr(\rho F)$$

It is also customary to assume, as we have already done implicitly in previous sections, that the marginal distributions associated with $W(a_k, b_\ell; \rho)$ should equal the quantum mechanical distributions associated with single measurements; i.e.,

$$\sum_{\ell} W(a_k, b_\ell; \rho) = Tr(\rho P_{\alpha_k}),$$

$$\sum_{k} W(a_k, b_\ell; \rho) = Tr(\rho P_{\beta_\ell}).$$

In ordinary classical statistics, where it is always possible to conceive of the measurements as revelations of possessed values, the possibility that this latter condition might be denied is almost unthinkable. However, in quantum theory where the bond linking observable and operator is far more subtle, where measurement results must be interpreted in the minimal way simply as numbers which emerge in response to measurement acts, it is not so unreasonable to suggest that the distribution and/or the values of $\mathcal{A}$-measurement results might depend on whether or not the $\mathcal{A}$-measurement is performed "alone" or in conjunction with other kinds of measurements. However, it is not difficult to see that such a dependence would raise serious conceptual problems which would greatly complicate the idea of measurement. Suppose, for example, that Th4 had to be narrowed to the statement that $M_1(\mathcal{A})$ yields elements of $\mathcal{E}(A)$, <u>provided</u> no physical operation except the $\mathcal{A}$-measurement itself is performed upon the system of interest. The empirical vagueness of such a proviso is immediately evident; supplementation of the quantal axioms by restrictions of this kind would require the

elucidation of criteria which distinguish physically between an $\mathcal{A}$-measurement "by itself" and an $\mathcal{A}$-measurement "in conjunction with other operations". Accordingly, for the sake of simplicity we assume, as is always done, that "an $\mathcal{A}$-measurement is an $\mathcal{A}$-measurement" regardless of what else happens concurrently and that marginal distributions are therefore always equal to those given by the quantum mechanical trace formula (Th1). Incidentally, as a consequence of this assumption, the uncertainty principle will apply to the results of simultaneous $\mathcal{A},\mathcal{B}$-measurements upon the members of a single ensemble (cf. section 2); this follows from the fact that the uncertainty theorem is derived from the ordinary quantal distributions for single observables which have just been assumed equal to the marginal distributions associated with simultaneous measurements.

There is a fair amount of literature dealing with the special case of (C_1) where $\mathcal{A}=\mathcal{X}$ (position) and $\mathcal{B}=\mathcal{P}$ (momentum), $[X,P]=i\hbar 1$. Because the spectra of X and P are continuous, an integral of a probability density $w(x,p;\psi)$ replaces the summation on probability W; for pure ρ i.e., $\rho=P_\psi$, (C_1) then becomes

$$\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dp\; w(x,p;\psi)\mathcal{F}(x,p)=\langle\psi,F\psi\rangle.$$

Now, <u>if</u> there were a general rule which provided F given $\mathcal{F}(x,p)$ it should then be possible to extract $w(x,p;\psi)$ from this equation by some inversion method. However, since we have found in previous sections that compound observables may have no operator counterpart at all, it would not be surprising if the $w(x,p;\psi)$ which satisfies (C_1) for a proposed correspondence $\mathcal{F}\leftrightarrow F$ turned out to be unsatisfactory.

Suppes[24] has argued along these lines that quantum mechanics cannot admit a proper distribution $w(x,p;\psi)$ and that simultaneous measurement of $\mathcal{X}$ and $\mathcal{P}$ must therefore be impossible. Although both conclusions are

false, it is instructive to examine his argument. The idea is to use (C_1) to obtain the characteristic function $\tilde{w}(\theta,\tau)$ for $w(x,p;\psi)$ and hence by Fourier inversion $w(x,p;\psi)$ itself. The characteristic function $\tilde{w}(\theta,\tau)$ is defined as the mean value of $F(x,p)=e^{i\theta x+i\tau p}$. Suppes assumes without special comment that if $\mathcal{X}\leftrightarrow X$, $\mathcal{P}\leftrightarrow P$, then

$$F(\mathcal{X},\mathcal{P})=e^{i\theta\mathcal{X}+i\tau\mathcal{P}}\longleftrightarrow e^{i\theta X+i\tau P}=F.$$

This suffices to determine $w(x,p;\psi)$ uniquely as the Fourier transform of $\langle\psi, e^{i\theta X+i\tau P}\psi\rangle$:

$$w(x,p;\psi)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\psi^*(x-\tfrac{u\hbar}{2})e^{-iup}\psi(x+\tfrac{u\hbar}{2})\,du.$$

Because this w is not positive semidefinite* yet is unique, Suppes concludes that $\mathcal{X}$ and $\mathcal{P}$ are incompatible observables.

Having seen the same erroneous conclusion drawn from (C_1) before (cf. section 4), it is not so difficult this time to reinterpret the mathematics. Suppes has not proved the sweeping assertion that $\mathcal{X}$ and $\mathcal{P}$ cannot be measured simultaneously; he has shown only that the correspondence

$$e^{i\theta\mathcal{X}+i\tau\mathcal{P}}\longleftrightarrow e^{i\theta X+i\tau P}$$

is not valid. We have seen earlier how von Neumann's incompatibility argument motivated the study of "quantum logics". Similarly, in a recent paper Suppes[25] deduces from several premises that "the functional or working logic of quantum mechanics is not classical". As might be expected, among these premises is his (false) conclusion that joint distributions in general do not exist in quantum theory. Hence this call for a quantum logic, like that discussed in previous sections, is inspired by the misinterpretation of a mathematical theorem.

The use of (C_1) on characteristic functions is not the only conceivable way to generate joint probabilities. Another scheme, developed by Margenau

---

*This w, known as the Wigner distribution,[26] is a valuable computational tool in statistical mechanics.

and Hill,[27] begins with the concept of covariance:

$$\mathrm{Cov}(a,B) = \langle aB\rangle - \langle a\rangle\langle B\rangle$$
$$= \sum_{kl} W(a_k, b_l; \psi)\, a_k b_l - \langle a\rangle\langle B\rangle.$$

The first problem is to find an Hermitean operator $F \leftrightarrow aB$ which leads to desirable properties for $\mathrm{Cov}(a,B)$. Once F is selected, the two covariance expressions above are equated to obtain what will be recognized as the (C_1) condition for $F(a,B) = aB$:

$$\sum_{kl} W(a_k, b_l; \psi)\, a_k b_l = \langle \psi, F\psi\rangle.$$

Because it gives the covariance reasonable properties, Margenau and Hill chose

$$F = \tfrac{1}{2}(AB + BA).$$

Unfortunately, the W's to which this leads are not true probabilities.

With our reinterpretation of von Neumann's "simultaneous measurability" theorem (cf. section 7), perhaps it is possible to understand why this covariance method was unsuccessful. Since that theorem, correctly interpreted, proves that the correspondence $aB \leftrightarrow \frac{1}{2}(AB+BA)$ is generally consistent only for commuting observables, it is inevitable that any theory built upon this correspondence will at some point break down. From this perspective, the work of Margenau and Hill must be regarded as further evidence against strong correspondence.

Although the elimination of strong correspondence does save quantum theory from self-contradiction, we now see that the consequent absence of operator-observable correspondences effectively leaves (C_1) useless for finding joint probabilities. It is therefore reasonable to shift the base of research from (C_1) to the marginal distribution requirements, for which no operators representing compound observables are needed. For a pure state $\rho = P_\psi$, we have

$$\sum_l W(a_k, b_l; \psi) = |\langle \alpha_k, \psi\rangle|^2,$$
$$\sum_k W(a_k, b_l; \psi) = |\langle \beta_l, \psi\rangle|^2.$$

There is at least one satisfactory solution to these equations, viz., $W_M(a_k, b_l; \psi) = |\langle \alpha_k, \psi \rangle|^2 |\langle \beta_l, \psi \rangle|^2$. As Margenau,[28] who proposed this distribution, has shown, the stochastically independent form of W_M in no way violates the spirit of quantum theory. (The uncertainty principle, for example, is derivable from W_M, a fact which shows again how illogical it is to interpret that principle in terms of mutual "interference" of measurements.) Nevertheless, it must be asked: are there other positive semidefinite W's which also satisfy the marginal distribution equations? Is W_M unique?

For the special case where $A = X, B = P, [X, P] = i\hbar 1$, L. Cohen[29] has laid the ground-work for answering this question by providing a canonical form for $w(x, p; \psi)$:

$$w(x,p;\psi) = \frac{1}{4\pi^2}\int d\theta \int du \int d\tau\, e^{-i\theta x - i\tau p + i\theta u} f(\theta,\tau)\, \psi^*(u - \tfrac{\tau\hbar}{2})\psi(u + \tfrac{\tau\hbar}{2});$$

any choice of $f(\theta, \tau)$ such that $f(\theta, 0) = f(0, \tau) = 1$ will generate a w satisfying the marginal requirements. The f for any given w is found by Fourier inversion. Unfortunately, no way has been found to isolate those f's which generate positive semidefinite w's.

However, if we restrict ourselves to operators with discrete spectra, the problem of finding <u>all</u> admissible $W(a_k, b_l; \psi)$ becomes more tractable. Consider the special case $\mathcal{A} = \frac{2}{\hbar}\mathcal{S}_z$, $\mathcal{B} = \frac{2}{\hbar}\mathcal{S}_x$, where $\mathcal{S}_x, \mathcal{S}_z$ are spin components for a spin-$\frac{1}{2}$ system. The operators which correspond to $\mathcal{A}$ and $\mathcal{B}$ are represented by the Pauli matrices:

$$\mathcal{A} \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = A, \quad \mathcal{B} \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = B.$$

If $\alpha_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\alpha_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\beta_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\beta_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $a_1 = 1, b_1 = 1, a_2 = -1, b_2 = -1$, these eigenvalue equations hold:

$$A\alpha_k = a_k \alpha_k, \quad B\beta_l = b_l \beta_l.$$

We seek all functions $W(a_k, b_\ell; \psi) \equiv W_{k\ell}(\psi)$ satisfying these conditions:

$$0 \le W_{k\ell}(\psi) \le 1 \,, \quad \sum_\ell W_{k\ell}(\psi) = |\langle \alpha_k, \psi \rangle|^2 \,, \quad \sum_k W_{k\ell}(\psi) = |\langle \beta_\ell, \psi \rangle|^2 .$$

For the present simple case, it is in fact possible to find all W's consistent with any given ψ. Consider, for example, $\psi = \begin{pmatrix} .9 \\ .1 \end{pmatrix}$. The marginal conditions are then

$$\begin{aligned} W_{11} + W_{12} &= .9 \\ W_{21} + W_{22} &= .1 \\ W_{11} + W_{21} &= .8 \\ W_{12} + W_{22} &= .2 \end{aligned}$$

Substituting $W_{22} = .2 - W_{12}$ into the second equation, we obtain three equations in three unknowns W_{11}, W_{12}, W_{21}:

$$\begin{aligned} W_{11} + W_{12} &= .9 \\ -W_{12} + W_{21} &= -.1 \\ W_{11} + W_{21} &= .8 \end{aligned}$$

An examination of the appropriate determinants reveals that these equations have an infinity of solutions. However, it is easy to find by using solid analytic geometry the straight line in (W_{11}, W_{12}, W_{21})-space which represents the set of real solutions. Then the condition $0 \le W_{k\ell} \le 1$ is used to delineate the segment of that line which represents <u>all</u> admissible sets (W_{11}, W_{12}, W_{21}) and W_{22} (via $W_{22} = .2 - W_{12}$). When this is done, there results a one parameter (z) family of positive semidefinite W's satisfying the marginal conditions:

$$W_z(a_k, b_\ell; \psi) = m_{k\ell} z + n_{k\ell} \,; \; z \in [0, 1],$$

$$(m_{k\ell}) = \begin{pmatrix} .1 & -.1 \\ -.1 & .1 \end{pmatrix} \,, \quad (n_{k\ell}) = \begin{pmatrix} .7 & .2 \\ .1 & 0 \end{pmatrix}.$$

(Although this was worked out in detail for a specific ψ, the same general procedure is of course applicable to any state vector in spinor sapce.)

One member of this family is the uncorrelated distribution $W_M(a_k, b_\ell; \psi) = |\langle a_k, \psi \rangle|^2 |\langle \beta_\ell, \psi \rangle|^2$; to find the corresponding

parameter τ_M consider the form $W_{\tau_M}(a_k, b_\ell) = f_k g_\ell$. From the equations, $f_1 g_1 = .1\tau + .7$, $f_2 g_1 = -.1\tau + .1$,

$f_1 g_2 = -.1\tau + .2$, $f_2 g_2 = .1\tau$,

two expressions for g_1/g_2 may be derived:

$$\frac{g_1}{g_2} = \frac{.1\tau + .7}{-.1\tau + .2} = \frac{-.1\tau + .1}{.1\tau}$$

Solving for τ, we obtain $\tau_M = .2$. All other τ values in the interval $[0,1]$ lead to correlated joint probability functions.

The purpose of this simple spinor space example has been to show that the marginal distribution requirements alone are not sufficient to determine a unique functional of ψ, $W(a_k, b_\ell; \psi)$. In general, given any state ρ, we therefore anticipate the existence of many distribution functions marginally consistent with that state.

It is difficult to evaluate the significance of this apparent multiplicity of W's. There are two possibilities: (1) For every state there exists a unique $W(a_k, b_\ell; \rho)$ governing the statistics of $\mathcal{A}, \mathcal{B}$-measurement results for real physical systems; or (2) The quantum mechanical state ρ does not determine the joint distribution of $\mathcal{A}, \mathcal{B}$-measurement results.

Alternative (1) has always been assumed implicitly throughout discussions of quantal joint probability. It is deeply rooted in the fundamental quantum belief that the density operator ρ embodies all that can conceivably be said about the measurement statistics from the ensemble to which ρ refers. Moreover, since only ρ obeys a causal law, it is the quantal construct "closest" to the classical ideal of physical state. The natural assumption to make therefore is that ρ determines $W(a_k, b_\ell)$ just as it determines $W_{\mathcal{A}}(a_k)$ and $W_{\mathcal{B}}(b_\ell)$. If so, we should attempt to formulate some physical criterion which, when required in conjunction with

the marginal rules, extracts the one true distribution from the set of marginally satisfactory ones.

There is in fact some precedent to support this plan. In the case of commuting observables, it also turns out that the marginal requirements alone do not determine the joint distribution. Consider, for example, two commuting spinor space observables $A \leftrightarrow \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$, $B \leftrightarrow \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}$, $a_1 \neq a_2, b_1 \neq b_2$; whose common eigenvectors are $\alpha_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\alpha_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. For the state $\psi = \begin{pmatrix} \sqrt{3/4} \\ \sqrt{1/4} \end{pmatrix}$, we obtain these marginal distributions:

$$W_{11} + W_{12} = W_{11} + W_{21} = 3/4,$$
$$W_{21} + W_{22} = W_{12} + W_{22} = 1/4.$$

These equations are satisfied by the following one parameter family of W's:

$$W_{kl} = m_{kl} z + n_{kl}, \quad z \in [0, 1],$$

where $(m_{kl}) = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix}$, $(n_{kl}) = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 \end{pmatrix}$.

Nevertheless, we are able to select from among these, by invoking further criteria, the single correct joint distribution. By working with an auxiliary observable C through which A and B may be trivially jointly measured, we find that

$$W_{11} = \frac{3}{4}, \quad W_{22} = \frac{1}{4}, \quad W_{12} = W_{21} = 0.$$

I.e., $z = 1$ gives the correct W_{kl}; the others are physically meaningless. (The general derivation of joint distributions for commuting observables was discussed fully in section 3.)

If alternative (2) is correct, then for the case of noncommuting observables there can be no similar criterion to distill a "true" distribution from the set of marginally satisfactory ones, for the density operator alone would not determine W. This is <u>not</u> the same as denying the existence of W; since noncommuting observables are simultaneously measurable, of course W exists. What alternative (2) suggests is that W does not exist <u>as a functional of state</u>, that simultaneous measurements are in some

sense theoretical anomalies not treatable by any comprehensive theory resembling the ordinary quantum theory of single measurements.

At this stage there is no firm basis for choosing between these alternate interpretations of the fact that many proper W's satisfy all valid conditions placed upon them. Shortly we shall return to this problem; but first it will be advantageous to do a bit of roaming in the relatively unexplored (due to aforementioned taboos) realm of simultaneous measurement theory.

9. A Search for Ordinary Simultaneous Measurements

As we have seen, attempts to approach the study of quantum joint probabilities via more or less natural random variable techniques seem invariably to be thwarted at some stage by ignorance of, or perhaps even the nonexistence of, operators corresponding to compound observables. It seems desirable therefore to develop a method for examining simultaneous measurements which does not depend on unknown operator-observable correspondence rules. To do this, we return to the general ideas concerning quantum measurement which were reviewed in section 1. As explained there, quantum mechanics is essentially a theory about systems which do not possess attributes (i.e., values of observables). Instead quantum observables are related to physical systems only in the dispositional sense conveyed by the philosophical doctrines of complementarity[30] and/or latency.[31] The primitive classical notion of possession ("System S has A-value a_k") is superseded by the primitive quantal measurement construct M_1 ("If $M_1(A)$ is performed on system S, the value a_k will result with probability..."). Accordingly, just as a theoretical explanation of measurement processes in classical physics involved relations among possessed attributes, a quantum theory of measurement at best describes connections among the unanalyzable

M_1's. Statements of such connections and associated empirical procedures constitute M_2, the usual scientific concept of measurement, or measurement scheme (operational definition, Margenau's epistemic correspondence rule). These constructs were exemplified in section 5 by a discussion of the time-of-flight $M_2(P)$ for momentum measurement; typically, the theory of $M_2(P)$ established a connection between $M_1(X)$, already operationally defined, and $M_1(P)$, which thereby acquired a theoretically validated empirical definition itself.

Because every physical process--hence any measurement scheme, single or joint--has a quantum theoretical description, it seems reasonable that whatever the correct joint probabilities are, they should be <u>derivable</u> within the framework of a quantum theory of M_2. That is, if a given procedure $M_2(X,P)$ is to be regarded as a method for simultaneous measurement of X and P, the scheme must be certified by a theory establishing relations between $M_1(X), M_1(P)$, and whatever "direct meter readings" are used as the basis for inference of simultaneous $M_1(X)$ and $M_1(P)$-results; from this analysis it should be possible in principle to find the probability for the occurrence of those "meter readings" which imply any given pair of X- and P-values. This measurement theoretical approach to the joint probability problem bypasses the operator-observable correspondence difficulty which obstructed the methods reviewed earlier. (All this will be clarified below by explicit examples.)

To develop these ideas further we next distinguish two kinds of M_2-theories: (1) ordinary and (2) historical. This distinction will later turn out to have considerable bearing on the problem of compatibility.

(1) An <u>ordinary</u> M_2 begins with system S in an arbitrary* state ρ_{t_0} at

---

*I.e., <u>practically</u> arbitrary. Recall, e.g., that the time-of-flight $M_2(P)$ did require $\psi(x)$ to have compact support, but no specific functional form was demanded.

some specified time t_0 and demonstrates how some operation upon S eventually leads to numbers from which may be inferred M_1-results to be associated with S in state ρ_{t_0}. It is to be especially noted that the state of S before t_0 is completely irrelevant. (2) An historical M_2-theory also seeks to certify some operation as a bona fide supplier of numbers which can be meaningfully interpreted as M_1-results for S in state ρ_{t_0}. However, unlike the ordinary type, the historical M_2-theory cannot be worked out without detailed information concerning the structure of ρ_{t_0}. Such information might be deduced from facts about the past history of the system, e.g., its state at some earlier time $t_1 < t_0$ plus its physical environment between t_1 and t_0. An example of each type appeared in section 5: the ordinary time-of-flight $M_2(P)$ and the historical time-of-flight $M_2(X, P)$.

Physically, the ordinary M_2-theories have been of greatest interest because they represent the idea of measurement in its purest form as a process applicable to a system at any instant independently of its past. An auto speedometer registers the speed of the car at any time regardless of the past wanderings of the vehicle; meteorological instruments record today's weather conditions with indifference toward yesterday's. Similarly, in quantum mechanics the language of M_1's presupposes that measurements are performed upon systems in states which are simply given without details as to the method of preparation. Accordingly, M_2-schemes for single observables (or commuting sets of observables) have been of the ordinary type. It is therefore natural to seek an ordinary M_2-theory covering the simultaneous measurement of several noncommuting observables. However, in view of the fact that both examples of simultaneous measurement given in section 5--the time-of-flight $M_2(X, P)$ and the use of two systems already correlated at the time of interest--were of the historical type, there is no reason

to be optimistic about finding any ordinary theory for simultaneous measurement.

We shall now outline two fairly general procedures which, at the outset, seem to be altogether plausible methods for achieving ordinary simultaneous measurement of two noncommuting observables. In both cases theoretical obstacles will eventually arise, and this will be interpreted as evidence that quantum theory does perhaps forbid ordinary simultaneous measurements. Deeper reasons to anticipate such a theoretic prohibition will then be explored.

Let S be the system upon which $M_2(\mathcal{A},\mathcal{B})$ is to be performed. Suppose a second system M with an observable $\mathcal{U}$, $\mathcal{U}\theta_\ell = u_\ell \theta_\ell$ interacts with S in such a way that the following correlations result:

$$T_{\mathcal{A},\mathcal{B}}(\psi \otimes \chi_0) = \sum_{k\ell} C_{k\ell}(\psi)\, \alpha_k \otimes \theta_\ell \quad , \text{ with marginal conditions}$$

$$\sum_\ell |C_{k\ell}(\psi)|^2 = |\langle \alpha_k, \psi\rangle|^2 \text{ and } \sum_k |C_{k\ell}(\psi)|^2 = |\langle \beta_\ell, \psi\rangle|^2,$$

where $T_{\mathcal{A},\mathcal{B}}$ is the evolution operator for this $M_2(\mathcal{A},\mathcal{B})$-scheme and ψ and χ_0 are the initial state vectors of S and M. Since these equations imply that post-interaction $M_1(\mathcal{A})$-results occur with the same probability as they would have before interaction and that a post-interaction $M_1(\mathcal{U})$ would yield u_n with the same probability that a pre-interaction $M_1(\mathcal{B})$ would have yielded b_n, an ordinary $M_2(\mathcal{A},\mathcal{B})$ procedure yielding $M_1(\mathcal{A})$- and $M_1(\mathcal{B})$-results for the instant before interaction may be defined as follows: after interaction, measure $\mathcal{A}$ on S, $\mathcal{U}$ on M; if the resulting number pair is (a_k, u_n), the pair (a_k, b_n) is declared the result of simultaneous $M_1(\mathcal{A})$ and $M_1(\mathcal{B})$ for S in the state ψ, i.e., just before interaction with M. If a $T_{\mathcal{A},\mathcal{B}}$ exists which can accomplish these correlations, the derivation of the associated joint probability distribution is immediate:

$$W(a_k, b_\ell; \psi; T_{\mathcal{A},\mathcal{B}}) = |C_{k\ell}(\psi)|^2.$$

We have indicated a possible dependence of W upon $T_{a,\mathcal{B}}$, i.e., upon the method of measurement, since in the case of joint measurements there is no a priori assurance that the distribution would be independent of this factor. (For single measurements, we know axiomatically that probabilities depend only on the state.) The important point here is that if such a $T_{a,\mathcal{B}}$ exists, we have an example of an ordinary theory of $M_2(a,\mathcal{B})$, complete with joint probability distribution. Unfortunately, no such $T_{a,\mathcal{B}}$ does exist.

To see why, consider first $\psi = \alpha_n$. In this case

$$\sum_\ell |C_{k\ell}(\alpha_n)|^2 = |\langle \alpha_k, \alpha_n \rangle|^2 = \delta_{kn}.$$

Since $|C_{k\ell}|^2 > 0$, this implies that $C_{k\ell}(\alpha_n) = 0, k \neq n$; thus $C_{k\ell}(\alpha_n)$ has the form $C_{k\ell}(\alpha_n) = \delta_{kn}\, g_\ell(\alpha_n)$ and accordingly

$$\sum_k |C_{k\ell}(\alpha_n)|^2 = |g_\ell(\alpha_n)|^2 = |\langle \beta_\ell, \alpha_n \rangle|^2.$$

Hence $T_{a,\mathcal{B}}(\alpha_n \otimes \chi_0) = \sum_{k\ell} \delta_{kn}\, g_\ell(\alpha_n)\, \alpha_k \otimes \theta_\ell$.

Now, by superposition we can derive the transformation associated with any ψ subjected to this type of measurement. If $\psi = \sum_n \langle \alpha_n, \psi \rangle \alpha_n$,

$$\begin{aligned} T_{a,\mathcal{B}}(\psi \otimes \chi_0) &= \sum_{nk\ell} \langle \alpha_n, \psi \rangle\, \delta_{kn}\, g_\ell(\alpha_n)\, \alpha_k \otimes \theta_\ell \\ &= \sum_{k\ell} \langle \alpha_k, \psi \rangle\, g_\ell(\alpha_k)\, \alpha_k \otimes \theta_\ell, \end{aligned}$$

i.e., $C_{k\ell}(\psi) = g_\ell(\alpha_k) \langle \alpha_k, \psi \rangle$.

Checking this form against the two required marginal conditions, we find that one of them is not in general satisfied: to be sure,

$$\sum_\ell |C_{k\ell}(\psi)|^2 = \left(\sum_\ell |g_\ell(\alpha_k)|^2\right) |\langle \alpha_k, \psi \rangle|^2 = |\langle \alpha_k, \psi \rangle|^2,$$

but

$$\sum_k |C_{k\ell}(\psi)|^2 = |g_\ell(\alpha_k)|^2 \left(\sum_k |\langle \alpha_k, \psi \rangle|^2\right) = |g_\ell(\alpha_k)|^2 = |\langle \beta_\ell, \alpha_k \rangle|^2 \neq |\langle \beta_\ell, \psi \rangle|^2.$$

Hence the ordinary $M_2(a,\mathcal{B})$ here envisaged is theoretically impossible.

The same conclusion is obtained if we attempt to construct an ordinary $M_2(a,\mathcal{B})$ based upon interaction of the primary system $\underline{S}$ with $\underline{\text{two}}$ measuring

devices, $\underline{M}(\alpha)$ and $\underline{M}(\beta)$, which we might call α- and β-"meters". If χ_0, ξ_0 are the initial states for the meters, $\mathcal{U}, \mathcal{V}$ are observables belonging to $\underline{M}(\alpha), \underline{M}(\beta)$, respectively, $\mathcal{U}\theta_m = u_m\theta_m$, $\mathcal{V}\eta_n = v_n\eta_n$, and $\{\phi_\ell\}$ is an orthonormal set of vectors in the Hilbert space of $\underline{S}$, the following correlation scheme would describe an $\mathcal{M}_2(\alpha, \beta)$:

$$T_{\alpha,\beta}(\psi \otimes \chi_0 \otimes \xi_0) = \sum_{\ell mn} s_{\ell mn}(\psi)\phi_\ell \otimes \theta_m \otimes \eta_n \quad , \text{ with}$$

marginal conditions $\sum_{\ell n} |s_{\ell mn}(\psi)|^2 = |\langle \alpha_m, \psi\rangle|^2$

and $\sum_{\ell m} |s_{\ell mn}(\psi)|^2 = |\langle \beta_n, \psi\rangle|^2$.

$\mathcal{M}_2(\alpha, \beta)$ would consist simply of "reading" the two meters (measuring $\mathcal{U}$ and $\mathcal{V}$) after the interaction; if (u_m, v_n) results, (a_m, b_n) is declared the result of simultaneous $\mathcal{M}_1(\alpha)$ and $\mathcal{M}_1(\beta)$ upon $\underline{S}$ in state ψ, i.e., just before the interaction. Once again the joint distribution would be easily derived:

$$W(a_m, b_n; \psi; T_{\alpha,\beta}) = \sum_\ell |s_{\ell mn}(\psi)|^2.$$

Like the previous example, this $\mathcal{M}_2(\alpha, \beta)$ is impossible; thus non-commuting observables cannot be simultaneously measured merely by letting $\underline{S}$ interact with two meters and reading the two separate results. To prove this, note first that for $\psi = \alpha_k$,

$$T_{\alpha,\beta}(\alpha_k \otimes \chi_0 \otimes \xi_0) = \sum_{\ell mn} \delta_{mk} g_{\ell n}(\alpha_k)\phi_\ell \otimes \theta_m \otimes \eta_n$$

$$= \sum_{\ell n} g_{\ell n}(\alpha_k)\phi_\ell \otimes \theta_k \otimes \eta_n,$$

$$\sum_{\ell m} |\delta_{mk} g_{\ell n}(\alpha_k)|^2 = \sum_\ell |g_{\ell n}(\alpha_k)|^2 = |\langle \beta_n, \alpha_k\rangle|^2,$$

$$\sum_{\ell n} |\delta_{mk} g_{\ell n}(\alpha_k)|^2 = |\langle \alpha_m, \alpha_k\rangle|^2 = \delta_{mk}.$$

Now consider the superposition $\psi = \sum_k \langle \alpha_k, \psi\rangle \alpha_k$; this leads to

$$T_{\alpha,\beta}(\psi \otimes \chi_0 \otimes \xi_0) = \sum_{k\ell n} \langle \alpha_k, \psi\rangle g_{\ell n}(\alpha_k)\phi_\ell \otimes \theta_k \otimes \eta_n,$$

i.e., $s_{\ell kn} = g_{\ell n}(\alpha_k)\langle \alpha_k, \psi\rangle$.

But this $s_{\ell kn}$ does not satisfy the second marginal condition:

$$\sum_{\ell k} |s_{\ell kn}|^2 = \sum_{\ell k} |\langle \alpha_k, \psi\rangle|^2 |g_{\ell n}(\alpha_k)|^2 = \sum_k |\langle \alpha_k, \psi\rangle|^2 \sum_\ell |g_{\ell n}(\alpha_k)|^2$$

$$= \sum_k |\langle \alpha_k, \psi\rangle|^2 |\langle \beta_n, \alpha_k\rangle|^2 \neq \left|\sum_m \langle \alpha_k, \psi\rangle\langle \beta_n, \alpha_k\rangle\right|^2 = |\langle \beta_n, \psi\rangle|^2.$$

The foregoing examples were presented to motivate this proposition:

(O) Ordinary simultaneous measurement of noncommuting observables is is theoretically impossible.

Of course merely citing two unsuccessful attempts to develop an ordinary $M_2(A,B)$ does not prove this proposition, but for the first time in the present study, we have found a good reason to suspect that quantum theory may indeed place some restriction upon joint measurability. However, if so, the qualification will not be a sweeping mandate to the effect that $M_2(A,B)$ is generally impossible, since that common version was refuted in section 5 by counterexamples. Rather (O) would mean only this: given at time t_0 a system S of unknown history, it is impossible to devise an operation $M_2(A,B)$ which leads to numbers (a_k, b_ℓ) interpretable as $M_1(A)$- and $M_1(B)$-results for time t_0.

Let us attempt to prove (O) by assuming it to be false and then trying to deduce a contradiction with quantum theory. That is, we now hypothesize that a theoretically certified, ordinary $M_2(A,B)$ does exist in spite of our preceding difficulties in formulating one. This means that given a system S in state $\rho_S(t_0)$ at t_0, $M_2(A,B)$ may be performed and will yield (a_k, b_ℓ), the results of simultaneous $M_1(A), M_1(B)$ at t_0. Now, suppose that S and another system C with which S is not interacting are regarded as one composite system and that the state of the latter at t_0 is $\rho_{S+C}(t_0)$. The density operators $\rho_{S+C}(t_0)$ and $\rho_S(t_0)$ will then be related by $Tr_C \rho_{S+C}(t_0) = \rho_S(t_0)$, where Tr_C denotes the trace over C's Hilbert space.

Since S is not in interaction with C, the efficacy of $M_2(A,B)$ cannot be affected by any statistical correlations which may inhere in the composite state $\rho_{S+C}(t_0)$. In particular, it may be that $\rho_{S+C}(t_0)$ involves double correlations of the kind employed in the famous Einstein-Podolsky-

Rosen[32] discussion concerning the completeness of quantum mechanics. If $\mathcal{A}$, $\mathcal{B}$ are the $\underline{S}$-observables to be measured and $\mathcal{U}, \mathcal{V}$ are noncommuting $\underline{C}$-observables with respective eigenvector sets $\{\theta_m\}, \{\eta_n\}$, such double correlation may be expressed* as follows:

$$\rho_{\underline{S}+\underline{C}} = P_\Psi, \quad \Psi = \sum_{jk} c_{jk} \alpha_j \otimes \theta_k = \sum_l d_l \beta_l \otimes \eta_l .$$

The first form of Ψ implies that from an $\mathcal{M}_1(\mathcal{A})$-result one may conclude what result a concurrent $\mathcal{M}_1(\mathcal{U})$ would have yielded; similarly the second form shows that if $\mathcal{M}_1(\mathcal{B})$ is performed the result that a simultaneous $\mathcal{M}_1(\mathcal{V})$ would yield may be inferred. Hence if the procedure $\mathcal{M}_2(\mathcal{A},\mathcal{B})$ is applied to $\underline{S}$, the results (a_j, b_l) lead to the <u>prediction with certainty</u> that simultaneous $\mathcal{M}_1(\mathcal{U})$ and $\mathcal{M}_1(\mathcal{V})$ would have yielded (u_j, v_l) at t_0. But system $\underline{C}$ is not disturbed by this process; during $\mathcal{M}_2(\mathcal{A},\mathcal{B})$, $\underline{C}$ evolves from its state at t_0 just as it would have if $\mathcal{M}_2(\mathcal{A},\mathcal{B})$ had not been performed on $\underline{S}$. However, $\mathcal{M}_2(\mathcal{A},\mathcal{B})$ has provided a means for dividing the $\underline{C}$-ensemble into identifiable subensembles. In particular, consider the subensemble comprised of those members of the $\underline{C}$-ensemble for which $\mathcal{M}_1(\mathcal{U})$ was certain to yield u_K and $\mathcal{M}_1(\mathcal{V})$ was certain to yield v_L at t_0. This (u_K, v_L)-subensemble must have a density operator $\rho_{KL}(t_0)$ since P2 and Th1 assure the existence of a ρ for every preparable ensemble. From the definition of the (u_K, v_L)-subensemble, it is clear that $\rho_{KL}(t_0)$ must satisfy these conditions:

$$W_{\mathcal{U}}[u_K; \rho_{KL}(t_0)] = Tr[\rho_{KL}(t_0) P_{\theta_K}] = 1,$$
$$W_{\mathcal{V}}[v_L; \rho_{KL}(t_0)] = Tr[\rho_{KL}(t_0) P_{\eta_L}] = 1,$$

and by Th3, $Tr[\rho_{KL}(t_0)] = 1$. By using the matrix representation of these equations, it is easy to see that the only possible solution is

---

*For specific examples, see Ref. 32 or Ref. 4.

$$\rho_{KL}(t_0) = P_{\theta_K} = P_{\eta_L}.$$

But since $[U, V] \neq 0$, the sets $\{\theta_k\}$ and $\{\eta_l\}$ cannot be identical; it is therefore always possible to choose u_K, v_L such that $\theta_K \neq \eta_L$. This means that $\rho_{KL}(t_0)$ does not exist, hence that the proposed preparation scheme for the (u_K, v_L)-subensemble is impossible. Thus some assumption upon which that scheme was based must be false.

To be specific, we must examine two key assumptions: (a) there exists an $M_2(A,B)$ performable upon any artibrary $\rho_S(t_0)$, i.e., an ordinary $M_2(A,B)$; (b) for every $\rho_S(t_0)$, it is possible to find $\rho_{S+C}(t_0)$ which incorporates the Einstein-Podolsky-Rosen type double correlation. It turns out that (b) is false; double correlations cannot be generated from arbitrary $\rho_S(t_0)$. For example, consider the common case $\rho_S = P_\psi$ and seek a vector Ψ such that $Tr_C P_\Psi = P_\psi$. From von Neumann's theorems about states of composite systems, it follows that ρ_C has the form $\rho_C = P_\chi$ and hence $\Psi = \psi \otimes \chi$, a form which cannot embrace double correlations. That this is the case may be seen by expanding $\psi \otimes \chi$ in terms of $\{\alpha_k \otimes \theta_m\}$ and of $\{\beta_l \otimes \eta_n\}$ and seeking conditions under which both expansions exhibit correlations:

$$\Psi = \sum_{km} \langle \alpha_k, \psi \rangle \langle \theta_m, \chi \rangle \alpha_k \otimes \theta_m = \sum_{ln} \langle \beta_l, \psi \rangle \langle \eta_n, \chi \rangle \beta_l \otimes \eta_n.$$

Consider only the first expansion; it will correlate a- and u-measurements if $\langle \alpha_k, \psi \rangle \langle \theta_m, \chi \rangle = \delta_{km} \langle \alpha_k, \psi \rangle \langle \theta_k, \chi \rangle$, a condition which can only be fulfilled by the trivial cases, $\chi = 0$ or $\psi = 0$, and by the conditions $\chi = \theta_m, \psi = \alpha_m$. But similarly it follows from the second expansion that $\chi = \eta_m, \psi = \beta_m$. Hence the very common case $\rho_S = P_\psi$ counters the assumption (b) that doubly correlated states are always available. Thus our EPR-inspired demonstration does not prove (0) conclusively.

It does, however, prove this: (0') <u>if</u> a system is in any doubly

correlated state $\rho_S(t_0)$, then ordinary simultaneous measurement of the noncommuting observables involved in the correlation is impossible. In other words, if assumption (a) were true at all, it would have to be modified to read as follows: (a') there exists an $M_2(A,B)$ performable upon any $\rho_S(t_0)$ except those which are doubly correlated in A and B --a rather strange statement.

Although the foregoing considerations do not rigorously prove that ordinary $M_2(A,B)$ are generally impossible, it is difficult to allay the suspicion that (0) may indeed be correct. The above reported failure of very reasonable attempts to develop ordinary M_2-theories together with the proof just given of (0') certainly point in the direction of (0). Moreover, the fact that we have been unable to find even specific cases of ordinary $M_2(A,B)$'s provides inductive support for (0).

Finally, it is instructive to summarize in a concise way the theoretical basis of our "partial proof" of (0), i.e., our proof of (0'). As we have seen, a consistent quantum axiom set does not forbid simultaneous measurement in general; it does, however, entail important prohibitions regarding "simultaneous preparation". For example, it is strictly impossible to devise a method for preparing systems certain to yield given Q- and P-values upon measurement; such a procedure would, among other things, violate the uncertainty principle. In short, certain "simultaneous preparations" are inconceivable in quantum mechanics.

Any proposed operation which leads to a contradiction with these basic restrictions on "simultaneous preparation" must be regarded as physically impossible. In the double correlation argument above, the operation in question happened to be a certain kind of measurement. However, this does not refute our earlier admonition (sec. 1) that measurement and preparation must not be equated as is customary with proponents of wave packet reduction.

Measurement and preparation are generally distinct, but, being constructs within the same theory, they can of course be related through propositions such as (O).

10. Some Examples of Quantal Joint Distributions

As we have just seen, it may be that the class of ordinary measurements cannot be extended to include simultaneous measurement of noncommuting observables. Thus it is plausible that the orthodox principle of incompatibility, though incorrect as a general proposition, might be a valid assertion about ordinary measurement schemes. Since the latter have customarily been emphasized in scientific practice, (O), if correct, would explain to some extent why it is that the standard doctrine of incompatibility, which is not only erroneous itself but is also often "derived" from false premises such as the projection postulate, has been able to survive and indeed flourish in the physical and philosophical literature for decades. Because of its practical, and perhaps general, validity for ordinary measurements, the orthodox principle of incompatibility has become a quantum platitude, a creed whose words are occasionally repeated but whose content is essentially ignored.

However, even if it turned out that ordinary joint M_2's were nonexistent, this would not eliminate all simultaneous measurement schemes. There would still be interesting historical procedures to study. Accordingly, the remainder of this section will be devoted to derivations of the joint distributions associated with several joint M_2's of the historical type. Perhaps another suitable adjective to describe this somewhat anomalous class of M_2's would be accidental, for as close scrutiny will reveal, each of these simultaneous measurement methods requires the quantum state involved to exhibit rather extraordinary properties. However, even though

they are "historical, accidental, and extraordinary", these schemes do represent valid simultaneous measurements of noncommuting observables, and they do therefore legitimately refute the claims that (a) such measurements are impossible and (b) that the associated joint distributions do not exist.

(1) Time-of-flight $\mathcal{M}_2(\mathcal{Q},\mathcal{P})$: The theory behind this method was presented in section 5, where it was shown that, under certain conditions, an $\mathcal{Q}$-measurement yielding x implied that a simultaneous $\mathcal{P}$-measurement would have yielded $p=\frac{mx}{t}$ with theoretical error which can be made arbitrarily small (in contradiction to the standard misinterpretation of the uncertainty theorem). Obviously, the joint distribution associated with this $\mathcal{M}_2(\mathcal{Q},\mathcal{P})$ is given by the following probability density:

$$W(x,p;\psi_t)=|\psi_t(x)|^2\,\delta\left(p-\frac{mx}{t}\right).$$

(2) Electric field $\mathcal{M}_2(\mathcal{Q},\mathcal{P})$: Consider an electron (charge e) in an uniform external electric field E. If $\psi(x,0)$ is the initial electron state, the state $\psi(x,t)$ at time t is given by the following unitary transformation:

$$\psi(x,t)=\sqrt{\frac{m}{2\pi i\hbar t}}\exp\frac{i}{\hbar}\left[\frac{mx^2}{2t}+\frac{1}{2}ftx-\frac{f^2t^3}{24m}\right]\int_{-\infty}^{\infty}dx_1\exp\frac{i}{\hbar}\left[-\frac{mxx_1}{t}+\frac{mx_1^2}{2t}+\frac{ftx_1}{2}\right]\psi(x_1,0)$$

where $f=eE$.

From this transformation follow two probability "matchings" on which an $\mathcal{M}_2(\mathcal{Q},\mathcal{P})$ may be based:

(a) $W_{\mathcal{P}}[p\in(p_1,p_2);\psi_0]=W_{\mathcal{P}}[p\in(p_1+ft,p_2+ft;\psi_t)]$,

(b) $W_{\mathcal{P}}[p\in(p_1,p_2);\psi_0]=W_{\mathcal{Q}}\left[x\in\left(\frac{p_1t}{m}+\frac{ft^2}{2m},\frac{p_2t}{m}+\frac{ft^2}{2m}\right);\psi_t\right]$, as $t\to\infty$, provided $\psi(x,0)$ is nonzero only in some finite interval $(-x_0,x_0)$.

Equation (a) means that $\mathcal{M}_1(\mathcal{P})$ at $t=0$ yields p with the same probability that $\mathcal{M}_1(\mathcal{P})$ at $t>0$ yields $p+ft$. Equation (b) provides this $\mathcal{M}_2(\mathcal{P})$: to determine what $\mathcal{M}_1(\mathcal{P})$ at $t=0$ would have yielded, measure $\mathcal{Q}$ at sufficiently large t and use the result x to evaluate $p=\frac{mx}{t}-\frac{ft}{2}$, which then counts as the result of $\mathcal{M}_1(\mathcal{P})$ at $t=0$. By analogy to the

time-of-flight method, we have only to combine (a) and (b) to obtain an $\mathcal{M}_2(\mathcal{Q},\mathcal{P})$: given an electron (with initial state $\psi(x,0)$ of compact support) in a uniform electric field, wait a very long time and then measure $\mathcal{Q}$; if the result is x, it is inferred through (a) and (b) that a simultaneous $\mathcal{M}_1(\mathcal{P})$ would have yielded $p=\left(\frac{mx}{t}-\frac{ft}{2}\right)+ft=\frac{mx}{t}+\frac{ft}{2}$. The joint probability density associated with this method is accordingly

$$W(x,p;\psi_t)=|\psi_t(x)|^2\,\delta\left(p-\frac{mx}{t}-\frac{ft}{2}\right).$$

To prove (a), it is sufficient to show that the evolution operator for the electron in the electric field transforms an eigenvector belonging to eigenvalue p into an eigenvector belonging to eigenvalue $p+ft$. Thus, let $\psi(x,0)=(2\pi\hbar)^{-1}\exp\frac{ipx}{\hbar}$ and determine the corresponding $\psi(x,t)$:

$$\psi(x,t)=\left\{\frac{1}{\sqrt{2\pi\hbar}}\sqrt{\frac{m}{2\pi i\hbar t}}\exp\frac{i}{\hbar}\left[\frac{mx^2}{2t}+\frac{1}{2}ftx-\frac{f^2t^3}{24m}\right]\right\}$$
$$\times\int_{-\infty}^{\infty}dx_1\exp\frac{i}{\hbar}\left[-\frac{mxx_1}{t}+\frac{mx_1^2}{2t}+\frac{1}{2}ftx_1+px_1\right]$$
$$=\left\{\right\}\int_{-\infty}^{\infty}dx_1\exp\left[-x_1^2\left(\frac{m}{2i\hbar t}\right)+ix_1\left(-\frac{mx}{t}+\frac{1}{2}\frac{ft}{\hbar}+\frac{p}{\hbar}\right)\right]$$
$$=\left\{\right\}\sqrt{\frac{2\pi i\hbar t}{m}}\exp\left[-\frac{it}{2m\hbar}\left(\frac{1}{2}ft-\frac{mx}{t}+p\right)^2\right]$$
$$=\left[\frac{1}{\sqrt{2\pi\hbar}}\exp\frac{i}{\hbar}(p+ft)x\right]\left[\exp-\frac{it}{\hbar}\left(\frac{p^2}{2m}+\frac{ftp}{2m}+\frac{f^2t^2}{6m}\right)\right],$$

which will be recognized as the desired eigenvector (the second bracket is just a phase factor independent of x).

To prove (b) let $\psi(x,0)$ denote any wave function which is nonzero only within an interval $(-x_0,x_0)$ and consider the following probability function:

$$W_{\mathcal{Q}}\left[x\in\left(\frac{p_1t}{m}+\frac{ft^2}{2m},\frac{p_2t}{m}+\frac{ft^2}{2m}\right);\psi_t\right]=\int_{a=\frac{p_1t}{m}+\frac{ft^2}{2m}}^{b=\frac{p_2t}{m}+\frac{ft^2}{2m}}|\psi(x,t)|^2dx$$
$$=\int_a^b\frac{m}{2\pi\hbar t}\left|\int_{-\infty}^{\infty}\exp\frac{i}{\hbar}\left[-\frac{mxx_1}{t}+\frac{mx_1^2}{2t}+\frac{ftx_1}{2}\right]\psi(x_1,0)\,dx_1\right|^2dx.$$

Under the substitution $x = \frac{pt}{m} + \frac{ft^2}{2m}$, this becomes

$$\int_{p_1}^{p_2} \left| \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-\frac{ipx_1}{\hbar}} \psi(x_1,0)\, e^{\frac{imx_1^2}{2\hbar t}} dx_1 \right|^2 dp,$$

which, by the same argument given in section 5 in the time-of-flight case, for sufficiently large t becomes arbitrarily close to

$$\int_{p_1}^{p_2} \left| \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-\frac{ipx_1}{\hbar}} \psi(x_1,0)\, dx_1 \right|^2 dp = W_\rho[p \in (p_1,p_2); \psi_0].$$

(3) Correlated systems $M_2(A_1, B_1)$: This general method was discussed at the end of section 5. Since the state of $S_1 + S_2$ is P_Ψ, where $\Psi = \sum_k c_k \alpha_k^{(1)} \otimes \alpha_k^{(2)}$, the state of S_1 alone is $\rho_1 = Tr_2 P_\Psi = \sum_k |c_k|^2 P_{\alpha_k^{(1)}}$. Now, the $M_1(A_2)$ which determines the desired $M_1(A_1)$-result also provides a method for resolving the S_1-ensemble into its pure subensembles $P_{\alpha_k^{(1)}}$. (As a matter of fact, this is a case where even the naive version of the projection postulate could be used without error.) Thus the S_1-subensemble to which the $M_1(A_1)$-result $a_k^{(1)}$ is assigned has state $P_{\alpha_k^{(1)}}$, and $M_1(B_1)$ on that subensemble would yield $b_l^{(1)}$ with probability $Tr(P_{\alpha_k^{(1)}} P_{\beta_l^{(1)}}) = |\langle \beta_l^{(1)}, \alpha_k^{(1)} \rangle|^2$. Since the fraction $|c_k|^2$ of the S_1-ensemble would be in the subensemble $P_{\alpha_k^{(1)}}$ it follows that the joint probability associated with this M_2 is

$$W(a_k^{(1)}, b_l^{(1)}; \rho_1) = |c_k|^2 |\langle \beta_l^{(1)}, \alpha_k^{(1)} \rangle|^2 = Tr(\rho_1 P_{\alpha_k^{(1)}}) Tr(P_{\alpha_k^{(1)}} P_{\beta_l^{(1)}}).$$

To summarize, the foregoing joint measurement schemes led to these probability functions:

(1) $w(x,p;\psi_t) = |\psi_t(x)|^2 \delta(p - \frac{mx}{t})$, $t \to \infty$;

(2) $w(x,p;\psi_t) = |\psi_t(x)|^2 \delta(p - \frac{mx}{t} - \frac{ft}{2})$, $t \to \infty$;

(3) $W(a_k^{(1)}, b_l^{(1)}; \rho_1) = Tr(\rho_1 P_{\alpha_k^{(1)}}) Tr(P_{\alpha_k^{(1)}} P_{\beta_l^{(1)}})$.

Note that in each case the two observables are stochastically dependent.

(4) Magnetic field $M_2(y, v_y)$: Since in a magnetic field the

canonical momentum is gauge dependent, we here consider the problem of measuring the y-component of position and velocity, y and v_y. If an electron in a uniform magnetic field $\vec{B}=B\hat{k}$ is initially in the state $\Psi(x,y,z;0)$, the state at time t is given by[33]

$$\Psi(x,y,z;t)=\left(\frac{m}{2\pi i\hbar t}\right)^{\frac{3}{2}}\left(\frac{\omega t/2}{\sin \omega t/2}\right)$$
$$\times\int_{-\infty}^{\infty}dx_1\int_{-\infty}^{\infty}dy_1\int_{-\infty}^{\infty}dz_1\,\exp\frac{im}{2\hbar}\left\{\frac{(z-z_1)^2}{t}+\frac{\omega}{2}\left(\cot\frac{\omega t}{2}\right)\left[(x-x_1)^2+(y-y_1)^2\right]+\omega(x_1y-xy_1)\right\}\Psi(x_1,y_1,z_1;0),$$

where $\omega\equiv\frac{eB}{mc}$ and the gauge is so chosen that the vector potential is $\vec{A}=\frac{B}{2}(-y\hat{i}+x\hat{j})$. From this transformation the following probability equations may be derived:

(a) $W_{v_y}[v_y\,\varepsilon\,(v_1,v_2);\Psi_0]=W_{v_y}[v_y\,\varepsilon\,(-v_2,-v_1);\Psi_\tau]$

where $\tau\equiv\frac{\pi}{\omega}$.

(b) If $\Psi(x,y,z;0)=\delta(x)\psi(y)\phi(z)$, then

$$W_{v_y}[v_y\,\varepsilon\,(v_1,v_2);\Psi_0]=W_x\left[x\,\varepsilon\left(\frac{2v_1}{\omega},\frac{2v_2}{\omega}\right);\Psi_\tau\right].$$

Proceeding as in previous examples, we may define an $M_2(y,v_y)$ as follows: to measure y and v_y at time τ for an electron initially in the state $\delta(x)\psi(y)\phi(z)$, measure x at τ and use the result x to evaluate $v_y=-\frac{\omega x}{2}$; also measure y at τ to get result y. From (a) and (b) it follows that the pair (y,v_y) may be regarded as the result of simultaneous $M_1(y)$ and $M_1(v_y)$ at τ. The joint distribution of y and v_y will depend on that of y and x since v_y is determined from x.

To prove (a) and (b) and to find the desired joint distribution, it will be necessary first to obtain the $\Psi(x,y,z;\tau)$ corresponding to $\Psi(x,y,z;0)=\delta(x)\psi(y)\phi(z)$.* It is convenient to work with the

---

* $\psi(y)$ and $\phi(z)$ are assumed normalized. However, since $\Psi(x,y,z;0)$ contains also a Dirac δ, it is only δ-function normalized; thus $\Psi(x,y,z;\tau)$ is also δ-function normalized. As a result, probabilities computed from these wave functions will involve meaningless $\delta(0)$ factors which will be replaced by unity at the end of such calculations. (This replacement will be indicated by using the symbol "$=$" instead of $=$.)

expansion

$$\Psi(x,y,z;0) = \delta(x)\int_{-\infty}^{\infty}\phi(p)\,\frac{e^{\frac{ipy}{\hbar}}}{\sqrt{2\pi\hbar}}\,dp\;f(z),$$

where $\phi(p)=\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}e^{-\frac{ipy}{\hbar}}\psi(y)\,dy$ is the canonical momentum space wave function corresponding to $\psi(y)$.

$$\Psi(x,y,z;\tfrac{\pi}{\omega}) = \left(\frac{m\omega}{2\pi^2 i\hbar}\right)\left(\frac{\pi}{2}\right)\left[\sqrt{\frac{m\omega}{2\pi^2 i\hbar}}\int_{-\infty}^{\infty}\exp\frac{im\omega}{2\pi\hbar}(z-z_1)^2 f(z_1)\,dz_1\right]$$

$$\times\int_{-\infty}^{\infty}dx_1\int_{-\infty}^{\infty}dy_1\exp\frac{im}{2\hbar}[\omega(xy-xy_1)]\,\delta(x_1)\int_{-\infty}^{\infty}dp\;\phi(p)\,\frac{e^{-\frac{ipy_1}{\hbar}}}{\sqrt{2\pi\hbar}}$$

$$=\left(\frac{m\omega}{4\pi i\hbar}\right)\left[f_{\text{free}}(z,\tau)\right]\int_{-\infty}^{\infty}dp\,\frac{\phi(p)}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}dy_1\exp-\frac{i}{\hbar}\left[\frac{m\omega x}{2}+p\right]y_1$$

$$=\left(\frac{m\omega}{4\pi i\hbar}\right)f_{\text{free}}(z,\tau)\int_{-\infty}^{\infty}dp\,\frac{\phi(p)}{\sqrt{2\pi\hbar}}\,2\pi\,\delta\left[\frac{1}{\hbar}\left(\frac{m\omega x}{2}+p\right)\right]$$

$$=\left[\sqrt{\frac{m\omega}{2}}\,\phi\left(\frac{m\omega x}{2}\right)\right]\left(\sqrt{\frac{m\omega}{4\pi\hbar}}\right)f_{\text{free}}(z,\tau).$$

(The factor in brackets is normalized.)

From this expression we next obtain

$$W_x\left[x\in\left(\tfrac{2\nu_1}{\omega},\tfrac{2\nu_2}{\omega}\right);\Psi_\tau\right]$$

$$=\int_{\frac{2\nu_1}{\omega}}^{\frac{2\nu_2}{\omega}}\left(\frac{m\omega}{2}\right)\left|\phi\left(\frac{m\omega x}{2}\right)\right|^2dx\int_{-\infty}^{\infty}\frac{m\omega}{4\pi\hbar}\,dy\int_{-\infty}^{\infty}\left|f_{\text{free}}(z,\tau)\right|^2dz$$

$$=\int_{\nu_1}^{\nu_2}\left|\phi(m\nu)\right|^2 m\,d\nu\;\delta(0)\,(1)$$

$$\text{“}=\text{”}\int_{\nu_1}^{\nu_2}\left|\phi(m\nu)\right|^2 m\,d\nu.$$

which must be compared to $W_{\nu_y}[\nu\in(\nu_1,\nu_2);\Psi_0]$, where

$$\Psi(x,y,z;0)=\delta(x)\,\psi(y)\,f(z).$$

The operator V_y such that $V_y \longleftrightarrow v_y$ is

$$V_{y_0} = \frac{1}{m} P_{y_0} - \frac{e}{mc} A_{y_0} = \frac{\hbar}{im} \frac{\partial}{\partial y} - \frac{\omega}{2} x ;$$

its eigenfunctions are $\{\eta_{x_0 p_0 z_0} = \delta(x-x_0) \frac{e^{\frac{i}{\hbar} y p_{y_0}}}{\sqrt{2\pi\hbar}} \delta(z-z_0)\}$. The V_y-eigenvalue corresponding to $\eta_{x_0 p_0 z_0}$ is $\left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)$. W_{v_y} is then calculated as follows:

$$W_{v_y}[\nu \varepsilon (\nu_1, \nu_2); \Psi_0] = \int_{\nu_1}^{\nu_2} d\nu \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \left| \langle \eta_{x_0 p_{y_0} z_0} \Psi_0 \rangle \right|^2 \delta\left[\nu - \left(\frac{p_{y_0}}{m} - \frac{\omega x_0}{2}\right)\right]$$

$$= \int_{\nu_1}^{\nu_2} d\nu \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \left| \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz\, \delta(x-x_0) \frac{e^{-\frac{i}{\hbar} y p_{y_0}}}{\sqrt{2\pi\hbar}} \delta(z-z_0) \delta(x) \psi(y) \delta(z) \right|^2$$

$$\times \delta\left[\nu - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right]$$

$$= \int_{\nu_1}^{\nu_2} d\nu \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0}\, \delta\left[\nu - \left(\frac{p_{y_0}}{m} - \frac{\omega x_0}{2}\right)\right] \left| \delta(x_0) \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-\frac{i p_{y_0} y}{\hbar}} \psi(y)\, dy \right|^2$$

$$= \int_{\nu_1}^{\nu_2} d\nu \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0}\, \delta\left[\nu - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right] \delta(x_0) \delta(x_0) \left| \phi(p_{y_0}) \right|^2$$

$$= \int_{\nu_1}^{\nu_2} d\nu \int_{-\infty}^{\infty} dp_{y_0}\, \delta\left[\nu - \frac{p_{y_0}}{m}\right] \delta(0) \left| \phi(p_{y_0}) \right|^2$$

$$\text{“}=\text{”} \int_{\nu_1}^{\nu_2} \left| \phi(m\nu) \right|^2 m\, d\nu .$$

Comparison with the expression for $W_x[x \varepsilon (\frac{2\nu_1}{\omega}, \frac{2\nu_2}{\omega}; \Psi_2]$ shows that (b) is true.

We next prove (a) for the specific Ψ_0 involved in (b), since only that special case is required for the M_2 under consideration.

$$W_{2_y}[v_y \in (-v_2,-v_1);\mathcal{F}_2] = \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \left| \langle \eta_{x_0,p_{y_0},z_0}, \mathcal{F}_2 \rangle \right|^2 \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega x_0}{2}\right)\right]$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \, \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right]$$

$$\times \left| \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz \, \delta(x - x_0) \frac{e^{-\frac{i p_{y_0} y}{\hbar}}}{\sqrt{2\pi\hbar}} \delta(z - z_0) \sqrt{\frac{m\omega}{2}} \phi\left(\frac{m\omega x}{2}\right) \sqrt{\frac{m\omega}{\pi\hbar}} \mathcal{F}_{\text{free}}(z,z) \right|^2$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \left(\frac{m\omega}{2}\right) \delta(p_{y_0}) \delta(p_{y_0}) \left| \phi\left(\frac{m\omega x_0}{2}\right) \right|^2 \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right]$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \, \frac{m\omega}{2} \left| \phi\left(\frac{m\omega x_0}{2}\right) \right|^2 \delta(0) \, \delta\left[v + \frac{\omega}{2} x_0\right]$$

$$= \int_{-v_2}^{-v_1} \left| \phi(-mv) \right|^2 m \, dv \; \delta(0)$$

$$\text{“}=\text{”} \int_{-v_2}^{-v_1} \left| \phi(-mv) \right|^2 m \, dv$$

$$= \int_{v_1}^{v_2} \left| \phi(mv) \right|^2 m \, dv .$$

This equals $W_{0_y}[v_y \in (v_1, v_2); \mathcal{F}_0]$; hence (a) is valid.

The joint probability density $w(y, v_y; \mathcal{F}_2)$ for this $M_2(y, v_y)$ is easily obtained. Since $M_1(v_y)$-results depend on $M_1(x)$-results, $w(y, v_y; \mathcal{F}_2)$ is simply related to $w_{x,y}(x, y; \mathcal{F}_2)$, which is calculable in the standard way:

$$w_{x,y}(x,y;\Psi_\tau)=\int_{-\infty}^{\infty}|\Psi(x,y,z;\tau)|^2dz=\left(\frac{mw}{4\pi\hbar}\right)\left|\sqrt{\frac{mw}{2}}\,\phi\left(\frac{mwx}{2}\right)\right|^2.$$

Since $v_y=-\frac{wx}{2}$ at τ, the density $w(y,v_y;\Psi_\tau)$ is given by

$$w(y,v_y;\Psi_\tau)=\frac{2}{w}w_{x,y}\left(-\frac{2v_y}{w},y;\Psi_\tau\right)=\left(\frac{mw}{4\pi\hbar}\right)m\left|\phi(-mv_y)\right|^2.$$

The fact (counterintuitive for some physicists) that simultaneous measurements of noncommuting observables could be stochastically independent and be in harmony with the basic structure of quantum mechanics has been emphasized by Margenau.[28] The above expression for $w(y,v_y;\Psi_\tau)$ shows that such independence of $M_1(y)$ and $M_1(v_y)$ actually obtains for this particular $M_2(y,v_y)$. To see this more clearly, note that $w(y,v_y;\Psi_\tau)=w_y(y)\,w_{v_y}(v_y)$, where w_y and w_{v_y} are computed in the usual quantum mechanical manner:

$$w_y(y)=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dz\,|\Psi(x,y,z;\tau)|^2$$

$$=\left(\int_{-\infty}^{\infty}dx\left|\sqrt{\frac{mw}{2}}\,\phi\left(\frac{mwx}{2}\right)\right|^2\right)\left(\int_{-\infty}^{\infty}dz\,|f_{free}(z)|^2\right)\left(\frac{mw}{4\pi\hbar}\right)$$

$$=\frac{mw}{4\pi\hbar},$$

$$w_{v_y}(v_y)=\int_{-\infty}^{\infty}dx_0\int_{-\infty}^{\infty}dp_{y_0}\int_{-\infty}^{\infty}dz_0\left|\langle\eta_{x_0p_{y_0}z_0};\Psi_\tau\rangle\right|^2\delta\left[v_y-\left(\frac{p_{y_0}}{m}-\frac{wx_0}{2}\right)\right]$$

$$=\delta(0)\,m\left|\phi(-mv_y)\right|^2\text{ ``}=\text{''}\,m\left|\phi(-mv_y)\right|^2.$$

Thus we have a counterexample to the idea that noncommuting observables must somehow "interfere" with each other during measurement. Although $[y,v_y]\neq 0$, simultaneous $M_1(y)$ and $M_1(v_y)$ performed via the present M_2 yield <u>stochastically independent</u> results. But of course the uncertainty principle still holds; indeed here $\Delta y\,\Delta v_y\to\infty$, where the Δ's are just standard deviations related to collectives of simultaneous y,

V_y-measurements.

In conclusion, we now summarize the important results obtained above:

1. If quantum theory had to be based upon the common assertion that to every physical observable there corresponds an Hermitean operator, then the theory would be self-contradictory, for that axiom proclaims the existence of certain operator-observable correspondences which, as a consequence of the remaining axioms, cannot exist.

2. A byproduct of these contradictions is von Neumann's theorem on simultaneous measurability, which is the only seemingly logical foundation ever given for the orthodox principle of incompatibility of noncommuting observables.

3. Once the concept of measurement is properly understood, it is possible to give explicit examples of quantum theoretical schemes for the exact simultaneous measurement of noncommuting observables.*

4. The contradiction between von Neumann's theorem (2) and the counterexamples (3) is simply a reflection of the fundamental contradictions (1).

5. If quantum theory entails only a weak correspondence between operators and observables, i.e., if it is assumed only that Hermitean

---

*It is important to realize that we have used the term measurement exclusively to refer to theoretically ideal processes which yield sharply defined numbers. (In a full treatment of the general nature of measurement, Margenau[34] calls such procedures "successful" measurements.) Occasionally one sees discussions of so-called imperfect "measurements" which presumably yield both a number x and an interval of "uncertainty" Δx covering that number, a concept invariably used without any clear definition. Recently several authors[35] who tacitly accept the projection postulate and therefore believe that noncommuting observables cannot be simultaneously measured with exactitude have advanced simultaneous "measurement" theories wherein the term measurement refers to these (in our opinion ill-defined) imperfect "measurements". Thus the similarity between these theories and the present work runs no deeper than the sound of the titles.

operators represent observables but not that all observables have operator representatives, then the contradictions disappear; in particular, von Neumann's theorem no longer follows from the axioms.

6. Moreover, none of the basic theorems which form the practical core of quantum physics is affected by this axiomatic shift from strong to weak correspondence.

7. The structure of quantum theory does seem to resist the formulation of any ordinary theory of measurement to describe simultaneous measurement processes; nevertheless, quantum mechanically certified simultaneous measurement schemes of the historical kind can be developed.

8. Derivation and comparison of the joint probability distributions associated with several methods of simultaneous measurement reveals that quantal joint probabilities may be either correlated (e.g., the time-of-flight case) or uncorrelated (e.g., the magnetic deflection type). Indeed the diversity of form among the specific distributions studied would seem to indicate that quantal joint probabilities for noncommuting observables are probably not functionals of state alone but depend as well on "historical" factors concerning particular M_2's.

Selected References

1. N. Bohr in Albert Einstein: Philosopher-Scientist, P. Schilpp (Ed.), Tudor (1949), 201-241.

2. W. Heisenberg, The Physical Principles of Quantum Theory, Chicago (1930).

3. E.g.: F. Mandl, Quantum Mechanics, Academic Press (1957); P. Roman, Advanced Quantum Theory, Addison-Wesley (1965).

4. J. L. Park, "The Nature of Quantum States" (Part I of Thesis).

5. J. L. Park, "Quantum Theoretical Concepts of Measurement" (Part II of Thesis).

6. J. von Neumann, Mathematical Foundations of Quantum Mechanics, Eng. trans. by R. T. Beyer, Princeton (1955), 351.

7. H. Margenau, Philo. Sci., 30 (1963), 6.

8. Ref. 5, sec. 10.

9. H. Margenau, Philo. Sci., 4 (1937), 352-356.

10. von Neumann, op. cit.

11. G. C. Wick, A. S. Wightman, and E. P. Wigner, Phys. Rev. 88 (1952), 101.

12. Ibid., 313-316; cf. also A. M. Gleason, J. Math. Mech., 6 (1957), 885.

13. E.g.: H. J. Groenewold, Physica, 12 (1946), 405; J. R. Shewell, Am. J. Phys., 27 (1959), 16.

14. J.v.Neumann, Ann. Math., 32 (1931).

15. Ref. 6, 225-230.

16. E.g.: G. W. Mackey, Mathematical Foundations of Quantum Mechanics, Benjamin (1963), 70; I. E. Segal, Ann. Math., 48 (1947), 930-948; V. S. Varadarajan, Comm. Pure Appl. Math., 15 (1962), 189-217.

17. E.g.: G. Birkhoff and J. von Neumann, Ann. Math., 37 (1936), 823; H. Reichenbach, Philosophic Foundations of Quantum Mechanics, U. of Calif. (1944); C. Piron, Helvetica Physica Acta, 37 (1964), 439.

18. R. P. Feynman, Quantum Mechanics and Path Integrals, McGraw-Hill (1965), 96-98.

19. Ref. 2, 20.

20. Ibid.

21. G. Temple, Nature, 135 (1935), 957; cf. also R. Peierls, Nature, 135 (1935), 957.

22. E.g.: F. Mandl, Introduction to Quantum Field Theory, Interscience (1959); R. F. Streater and A. S. Wightman, PCT, Spin and Statistics and All That, Benjamin (1964), 100.

23. E.g.: P. A. M. Dirac, Rev. Mod. Phys., 17 (1945), 195; J. E. Moyal, Proc. Camb. Phil. Soc., 45 (1949), 99.

24. P. Suppes, Philo. Sci., 28 (1961), 378 and in Philosophy of Science: The Delaware Seminar, v.2.

25. P. Suppes, Phil. Sci., 33 (1966), 14-21.

26. E. Wigner, Phys. Rev., 40 (1932), 749.

27. H. Margenau and R. N. Hill, Progr. Theor. Phys., 26 (1961), 727.

28. H. Margenau, Ann. Phys. (USA), 23 (1963), 469.

29. L. Cohen, J. Math. Phys., 7 (1966), 781.

30. N. Bohr, Essays 1958-1962 on Atomic Physics and Human Knowledge, Interscience (1963), 1-7.

31. H. Margenau, The Nature of Physical Reality, McGraw-Hill (1950), 171-177.

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

33. Ref. 18, 64.

34. Ref. 31, 369-375.

35. E. Arthurs and J. L. Kelly, Jr., Bell Sys. Tech. J. 44 (1965), 725; C. Y. She and H. Heffner, Phys. Rev., 152 (1966), 1103; E. Prugovečki, "On a Theory of Measurement of Incompatible Observables in Quantum Mechanics" (to be published).