The MIT Press
Cambridge, Massachusetts, and
London, England

Perspectives in Quantum Theory

Essays in Honor of Alfred Landé

edited by
Wolfgang Yourgrau
and
Alwyn van der Merwe

Five

James L. Park and
Henry Margenau

The Logic of Noncommutability of Quantum-Mechanical Operators—and Its Empirical Consequences

We wish to honor Alfred Landé by scrutinizing in this article one of the shibboleths of the quantum doctrine: the impossibility of performing simultaneous measurements of noncommuting observables. In his book[1] Landé regards as a half-truth the proposition: p and q cannot be measured simultaneously. The present paper presents an examination and indeed a justification of this claim.

1. The Compatibility Problem

Quantum physics, in using *operators* instead of functions to represent observables, complicates the relationship between its observables (i.e. their mathematical representatives) and the empirical numbers to which they must ultimately refer. Perhaps the most controversial difficulty associated with this operator-observable correspondence arises from the commutative law of arithmetic, viz. if a and b are numbers, then $ab = ba$. Naturally this law applies to all measurement results, quite independently of the theory by which they are interpreted. In quantum theory, however, the associated pairs of Hermitian operators do not necessarily commute.

Understandably, the presence in quantum theory of noncommuting observables has from the beginning elicited a great deal of curiosity. Some kind of physical interpretation must be given; the fact, for instance, 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*, that is, it is impossible to perform upon a single system *simultaneous measurements* of two such observables. The present paper is an abbreviated report of a systematic analysis of this famous principle of impotence; but first, as a prelude to the substance of the article, we feel it appropriate to review briefly the more common—and frequently illogical—arguments typically advanced in behalf of the doctrine in question. Many of them have already been subjected to criticism by Landé.

The typical historical account of quantum theory from Planck to the present endeavors to present a rather smooth transition from the "old quantum theory" (Bohr atom, particulate photon, classical ontology) to the "new quantum theory" (state vectors, probability, complementarity). Yet any discussion about modern quantum theory which employs concepts peculiar to the "old" to demonstrate alleged features of the "new" is of little value. The following sections therefore contain a *logical* study of the notion of compatibility entirely within the axiomatic framework of (new!) quantum theory, independently of whatever dreams, intuitions, or *Gedankenexperimente* historically might have inspired its ingenious creators.

(1) UNCERTAINTY PRINCIPLE. Many *Gedankenexperimente* have been designed to illustrate Heisenberg's famous law; unfortunately, the false impression is often conveyed that this principle, which is actually a theorem about standard deviations in collectives of measurement results, imposes restrictions on *measurability*. Simple common-sense arguments quite unrelated to the quantum theory could easily be adduced to show the elementary absurdity of such an inference.

(2) PROJECTION POSTULATE (NAIVE VERSION).* Frequently appended to the useful postulates of quantum mechanics is one which, if it were correct, could lead to the incompatibility doctrine as a theorem. It is the notion of wave-packet reduction, according to which any 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. The reasoning is: If simul-

* The fundamental irrationality, together with the mathematical strangeness, of the view that a single observation shall in general fix the probability distribution (state vector) of an entire ensemble has been emphasized repeatedly by one of the present authors. (Refs. 4 and 5.) This point is further elaborated in Refs. 2 and 3.

taneous measurement of noncommuting observables were possible, such an act could leave a system in a nonexistent state. This argument is, however, unworthy of serious consideration, since the idea of wave-packet reduction does not survive close scrutiny.[2–5]

(3) 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 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 universal quantal axiom.

(4) 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 has been examined carefully in another paper by the present writers.[9] There this first probability argument in behalf of the incompatibility principle is traced to the same fundamental errors which underlie the following argument (5). Since arguments (4) and (5) stand or fall together, we shall here concentrate our attention on (5).

(5) VON NEUMANN'S SIMULTANEOUS MEASURABILITY THEOREM. In his classic work on quantum mechanics, von Neumann proved a theorem which provides 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)–(4), it is a logical deduction from a seemingly reasonable quantum axiom set which does not include the projection postulate.* (Cf.

* To be sure, the projection postulate does appear in von Neumann's book, but it plays no role in the theorem here considered.

Sec. 4.) The ensuing sections will emphasize argument (5), the only extant evidence for incompatibility which is firmly embedded in the basic mathematical structure of modern quantum theory.

Because (5) arises deep in the theoretical framework of quantum mechanics, it seems desirable here to interpolate a brief survey of basic quantum axiomatics in order to furnish a basis for distinguishing clearly which quantum statements are hypotheses and which ones are *derivable* propositions. Only in this way can the deduction in (5) be properly evaluated.

As everywhere else, the objects of study in quantum mechanics are called *physical systems*. Associated with them are the constructs known as *observables*, which in turn are correlated via rules of correspondence to empirical operations that generate numbers. These operations are *measurements*. The numbers they produce are called *measurement results*, and it is the function of quantum theory to regularize, interpret, and make predictions about them. Specifically, quantum physics deals with problems of this kind: Given an actual (in contradistinction to a *Gedanken*-type) repeatable laboratory procedure Π for the *preparation* of physical systems, what will be the statistical distribution of measurement results obtained from observations performed upon an ensemble of systems all prepared in accordance with Π? This question may refer to any observable and to measurements at any given time after preparation.

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 and indeed the common-sense 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. The possessive adherence of observables to systems fails. This peculiarity of quantum observables has been characterized by one of the authors as *latency*.[10] A brief explanation of the idea of latency relevant to the present problem is given in Ref. 9.

The possessed quality of all classical observables brought the ideas 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. In quantum theory, however, the constructs measurement and preparation must be disjoined. Failure to do so leads to an erroneous interpretation of the projection postulate.[11] The correlation between measured values and the state of a system is less direct, and thus the idea of measurement needs more careful analysis than is ordinarily necessary.

To measure the *position* of an object, one juxtaposes the object with a scale and identifies its position with the scale mark. No further analysis is required, because the tacit belief that the object *has* the position avoids every complication. In this simple sense, a position measurement is the establishment of a self-evident correspondence between a set of numbers and the values of an observable—self-explanatory and self-validating. This manner of correspondence between measured numbers and values of the observable is part of every measurement in classical physics as well as in quantum mechanics. It is often called "direct" measurement, and we shall designate it by $\mathcal{M}_1$.*

The measurement of *velocity* is a little less obvious. If the speed of a moving vehicle is to be determined, the reading of a speedometer can be noted. What this measurement delivers directly via the $\mathcal{M}_1$ concept is the position of the needle. To interpret this position as a velocity requires assumptions of an interpretative, theoretical sort, ideas beyond the content of the mere association $\mathcal{M}_1$. It involves a mathematical analysis of the instrument, which leads to proof of a further correlation between the possessed position values of the

* In this essay we employ script letters ($\mathcal{M}$, $\mathcal{A}$, $\mathcal{B}$, . . .) to designate observables [the general concept of measurement is regarded as an observable]; capital italic letters (A, B) denote operators and lowercase italic letters (a, b) numerical measured values; lowercase Greek letters (α, β), except ρ (the density operator), denote quantum state vectors.

needle and the (in this case possessed) speed values of the moving vehicle. The ingredient of the measurement concept which establishes this theoretical correlation will be denoted by $\mathcal{M}_2$. Thus $\mathcal{M}_2(\mathcal{A}, \mathcal{B}, \ldots)$ represents any empirical procedure yielding numbers $a, b, \ldots$ which *through a theory* can be interpreted as the values *associated* with observables $\mathcal{A}, \mathcal{B}, \ldots$.

The fundamental difference between classical and quantal physics lies in their respective conceptions of the nature of this *association* between the measured values of observables and the systems with which the observables are said to be associated. Consider again the speedometer. Classically, its operation was explained by an $\mathcal{M}_2$ theory, which established a correlation between the *possessed* position values of the needle and the *possessed* speed values of the vehicle. Quantum mechanically, a correlation cannot be so simply described, because the intrinsic *latency* of quantum observables disavows the classically implicit presupposition that the system *possesses* any speed value to be correlated with the needle position. What construct then plays the same role in quantum physics that possession does in classical physics?

An examination of quantum axiomatics reveals (cf. Ref. 3) that the strongest kind of correlation statement for the speedometer which can be conceived within the framework of quantum physics has this form: The joint probability that a *measurement* of the needle position will yield x_k and a simultaneous *measurement* of the speed of the vehicle will yield v_l vanishes unless $l = k$.

Thus in quantum mechanics the classical notion of *possession* is replaced by a primitive construct, commonly called *measurement*, which is not, however, the concept of measurement designated above as $\mathcal{M}_2$. Indeed, as just exemplified, an $\mathcal{M}_2$ cannot be described quantum theoretically, except in terms of this primitive measurement construct, which will henceforth be called $\mathcal{M}_1$ to indicate that it is an abstract generalization of the elementary act of "direct" observation which was earlier denoted by $\mathcal{M}_1$. Since quantum physics, owing to the characteristic latency of its observables, cannot make statements about correlations among possessed values of observables, it speaks instead of correlations among the results of primitive "direct" measurements $\mathcal{M}_1$, which are in practice *theoretical* constructs which can no more be performed "directly" than could the possessed values of, say, the speed of a classical body be perceived "directly."

For a microphysical illustration, consider the measurement of electron position using a scintillation screen. The theory ($\mathcal{M}_2$) which justifies identification of the location of the scintillation with the position of an electron establishes a correlation not between two possessed attributes (positions of electron and scintillation) but between two $\mathcal{M}_1$'s: one the genuinely "direct" observation of the scintillation, the other the more sophisticated, abstract, and theoretically primitive relation that the observable position bears to an electron in quantum physics. These $\mathcal{M}_1$'s are governed by certain basic axioms to which Sec. 2 will be devoted.

Using the foregoing concepts, it is possible to define precisely what is meant by simultaneous measurability of two observables: *Observables $\mathcal{A}$ and $\mathcal{B}$ will be termed compatible, simultaneously measurable, or jointly measurable if there exists an $\mathcal{M}_2(\mathcal{A}, \mathcal{B})$*, that is, an operation furnishing two numbers a, b with the same probabilities that quantum theory confers upon the two propositions "$\mathcal{M}_1(\mathcal{A})$ yields a" and "$\mathcal{M}_1(\mathcal{B})$ 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

The basic postulates of quantum physics will now be stated, and several important theorems will then be reviewed.

P1 CORRESPONDENCE POSTULATE

(Some) linear Hermitian operators on Hilbert space which have complete orthonormal sets of eigenvectors correspond to physical observables. If operator A corresponds to observable $\mathscr{A}$, then the operation $\mathscr{F}(A)$, where $\mathscr{F}$ is a function, corresponds to observable $\mathscr{F}(\mathscr{A})$.

We shall use the symbol $\leftrightarrow$ to represent this operator-observable correspondence; thus $A \leftrightarrow \mathscr{A}$ means A "corresponds to" $\mathscr{A}$ in the sense of P1. The observable $\mathscr{F}(\mathscr{A})$ is defined operationally as follows: Measure $\mathscr{A}$ and use the result a to evaluate the given function $\mathscr{F}$; the number $\mathscr{F}(a)$ is then the result of an $\mathscr{F}(\mathscr{A})$-measurement. The function $\mathscr{F}$ of operator A, $\mathscr{F}(A)$, 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 $\mathscr{F}(A)$ is then simply

$$\sum_k \mathscr{F}(a_k) P_{\alpha_k}.$$

P2 MEAN VALUE POSTULATE

To every ensemble of identically prepared systems there corresponds a real linear functional of the Hermitian operators, $m(A)$, such that if $A \leftrightarrow \mathscr{A}$, the value of $m(A)$ is the arithmetic mean $\langle \mathscr{A} \rangle$ of the results of $\mathscr{A}$-measurements* performed on the member systems of the ensemble.

The content of P1 and P2 is slightly different from von Neumann's axiomatizations. In the original form of the Correspondence Postulate, observables and Hermitian operators were assumed to stand in one-to-one correspondence; the postulate included both of the following statements:

(1) Every observable has an Hermitian operator representative.

(2) Every Hermitian operator corresponds to a physical observable.

* Reference is here to the primary quantum-measurement construct $\mathscr{M}_1(\mathscr{A})$.

In 1952, Wick, Wightman, and Wigner[12] effectively challenged the symmetry of this quantal correspondence by introducing the concept of superselection rules, i.e. assertions which declare certain Hermitian operators to be unobservable in principle. To embrace superselection rules with minimal theoretic change, the word *every* in (2) is replaced by *some*:

(2′) Some Hermitian operators correspond to physical observables.

Just as superselection rules challenge the word *every* in (2), an important facet of the compatibility problem hinges on the word *every* in proposition (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* correspondence means that both (2′) *and* (1) are assumed; *weak* correspondence means that the Correspondence Postulate includes (2′) but *not* (1), as in P1.

In subsequent sections, the relationship of this choice of correspondence schemes to the problem of compatibility will be developed, and we shall demonstrate that only the weak type (P1) is acceptable. Henceforth we distinguish between P1S—von Neumann, strong—and our P1.

Several "elementary" quantum theorems will now be stated without proof. Although the content of these theorems is well known, the fact is not always acknowledged that they are *theorems*, i.e. derivable from P1 and P2: P1 and P2 (or their equivalents) rigorously *imply* all the general propositions of quantum statics.

TH. 1[13]

For each mean value functional $m(A)$ there exists an Hermitian operator ρ such that for each A,

$$m(A) = \mathrm{Tr}\,(\rho A).$$

The Hermitian operator ρ, known as the statistical operator or density operator, is not only an "index" of measurement statistics but also the seat of causality in quantum physics. For this reason, we shall call ρ the quantum *state* 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 like systems is given by

$$\rho(t_2) = T(t_2, t_1)\rho(t_1)T^{\dagger}(t_2, t_1).$$

In the following theorems, we assume the Hermitian operators to have discrete spectra; but similar propositions hold for the continuous case too.

TH. 2

The probability $W_{\mathscr{A}}(a_k; \rho)$ that an $\mathscr{A}$-measurement on a system from an ensemble with density operator ρ will yield the A-eigenvalue a_k is given by

$$W_{\mathscr{A}}(a_k; \rho) = \mathrm{Tr}\ (\rho P_{\mathscr{H}_k}),$$

where $\mathscr{H}_k$ is the subspace belonging to a_k.

TH. 3

$$\mathrm{Tr}\ \rho = 1.$$

TH. 4

The only possible results of $\mathscr{A}$-measurements are the eigenvalues $\{a_k\}$ of A, where $A \leftrightarrow \mathscr{A}$.

TH. 5

The density operator ρ is positive semidefinite.

Careful analysis (cf. Ref. 9) of the foregoing theorems reveals that only *weak* correspondence need be invoked to prove them; they would of course still follow, however, if P1 were replaced by an axiom of strong correspondence, namely

P1S
The set of physical observables is in one-to-one correspondence with the set of linear Hermitian operators on Hilbert space with complete orthonormal sets of eigenvectors. Thus, if $A \leftrightarrow \mathscr{A}$, then $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$.

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 via $\mathscr{M}_1(\mathscr{A})$. Indeed, the absence of a similar joint measurement construct $\mathscr{M}_1(\mathscr{A}, \mathscr{B}, \ldots)$ appears to justify the conclusion that quantum theory is noncommittal 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. This indifference is illusory, for P1S and P2 do in fact place severe restrictions upon simultaneous measurements through a theorem to be reviewed in Sec. 4.

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 $\mathscr{F}(\mathscr{A}, \mathscr{B})$ may be operationally defined as follows: Measure $\mathscr{A}$ and $\mathscr{B}$ simultaneously, and substitute the results a and b into the function $\mathscr{F}(a, b)$; the value $f = \mathscr{F}(a, b)$ is then the result of the $\mathscr{F}(\mathscr{A}, \mathscr{B})$-measurement. Then by P1S, there exists an operator F to represent $\mathscr{F}(\mathscr{A}, \mathscr{B})$; hence if F is known, $\mathscr{F}(\mathscr{A}, \mathscr{B})$-measurements are subject to quantum-mechanical analysis, and in this sense joint measurements would be in the domain of the ordinary quantum theory of $\mathscr{M}_1$'s.

This leads to an old quantum problem.[14] Given the correspondences $A \leftrightarrow \mathscr{A}$, $B \leftrightarrow \mathscr{B}, \ldots$ and a compound observable $\mathscr{F}(\mathscr{A}, \mathscr{B}, \ldots)$, what F corresponds to $\mathscr{F}$? If P1S is adopted, the existence of such an F is assured (provided, of course, that $\mathscr{A}, \mathscr{B}, \ldots$ are simultaneously measurable). If, however, only the weaker P1 holds, the existence of an F such that $F \leftrightarrow \mathscr{F}(\mathscr{A}, \mathscr{B}, \ldots)$ is by no means guaranteed. In neither case is there a general prescription for finding F, to be sure; but it is obviously necessary that all deductions based on a proposed F be consistent with P2, the definition of $\mathscr{F}$, and the theorems reviewed above. Th. 1 and Th. 4 particularly suggest useful consistency conditions. To formulate them, we employ the following notation.

Define the sets $\mathscr{E}(A)$ and $\mathscr{N}(\mathscr{F})$ thus: $\mathscr{E}(A)$ comprises the eigenvalues belonging to the operator A; $\mathscr{N}(\mathscr{F})$ is the set of obtainable measurement results associated with an observable $\mathscr{F}$. When $\mathscr{F} = \mathscr{A}$, then $\mathscr{N}(\mathscr{A}) = \mathscr{E}(A)$ by Th. 4. However, when $\mathscr{F}$ is a function, let us say of $\mathscr{A}$ and $\mathscr{B}$, then it is possible that correlations between $\mathscr{A}$ and $\mathscr{B}$ might preclude the occurrence of certain a priori conceivable values of $\mathscr{F}$, i.e. preclude certain of the values $\mathscr{F}(a_k, b_l)$ calculable from eigenvalues of A and B under the a priori assumption that all eigenvalue pairs (a_k, b_l) are possible. In this contingency, $\mathscr{E}(F) \subset \mathscr{N}(\mathscr{F})$. Finally, for a state ρ, let $W(a_k, b_l, \ldots; \rho)$ denote the joint probability that simultaneous $\mathscr{A}, \mathscr{B}, \ldots$-measurements yield $a_k, b_l, \ldots$.

Two consistency conditions may then be expressed in this manner: If $F \leftrightarrow \mathscr{F}(\mathscr{A}, \mathscr{B}, \ldots)$, then

$$(\mathrm{C}_1) \quad \sum_{kl} W(a_k, b_l, \ldots; \rho)\mathscr{F}(a_k, b_l, \ldots) = \mathrm{Tr}\,(\rho F),$$

for every ρ, and

$$(\mathrm{C}_2) \quad \mathscr{E}(F) \subseteq \mathscr{N}[\mathscr{F}(\mathscr{A}, \mathscr{B}, \ldots)].$$

Condition (C_1) arises from Th. 1 and the definition of $\mathscr{F}$, while (C_2) is needed to prevent conflict with Th. 4. The usefulness of (C_1) must however be questioned, for the joint probability W is so far unknown. Nevertheless, this condition is not undiscriminating, since for the proper choice of $\mathscr{F}$ it becomes independent of the form of W. (Cf. Ref. 9.)

Both P1 and P1S involve explicit postulation of the correspondence $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$, and a survey[15] of the proofs of Theorems 1–5 indicates clearly the value of that rule. Since Th. 2 (i.e. the form of $W_{\mathscr{A}}$) is the cornerstone of practical calculations in quantum theory and therefore not a proposition that could easily be challenged, the following theorem suggests very strongly that the correspondence $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$ could not reasonably be removed from the quantum axiom set.

CONSISTENCY THEOREM

If $W_{\mathscr{A}}(a_k; \rho) = \mathrm{Tr}\,(\rho P_{\mathscr{H}_k})$ and if there exists an operator F such that $F \leftrightarrow \mathscr{F}(\mathscr{A})$, then $F = \mathscr{F}(A)$, where $A \leftrightarrow \mathscr{A}$. The proof is simple.

The operator F must satisfy consistency condition (C_1):

$$\sum_k \mathrm{Tr}\,(\rho P_{\mathscr{H}_k})\mathscr{F}(a_k) = \mathrm{Tr}\,(\rho F).$$

Thus,

$$\mathrm{Tr}\,[\rho(F - \sum_k \mathscr{F}(a_k)P_{\mathscr{H}_k})] = 0 \qquad \text{for every } \rho,$$

which implies

$$F = \sum_k \mathscr{F}(a_k)P_{\mathscr{H}_k} = \mathscr{F}(A).$$

This result also satisfies (C_2).

The special case of $\mathscr{F}(\mathscr{A})$ that has received most attention is a fairly complicated one: $\mathscr{F}(\mathscr{A}) = (\mathscr{A} - \langle\mathscr{A}\rangle)^2$, where $\langle\mathscr{A}\rangle$ is a real constant which is the arithmetic mean of the $\mathscr{M}_1(\mathscr{A})$'s on the ensemble of interest. Using P2, we see that $\mathscr{F}(\mathscr{A}) = (\mathscr{A} - m(A))^2$; then, by the correspondence rule in P1, $\mathscr{F}(\mathscr{A}) \leftrightarrow (A - m(A)1)^2$. By definition,

$$(\Delta\mathscr{A})^2 = m[(A - m(A)1)^2] = \langle(\mathscr{A} - \langle\mathscr{A}\rangle)^2\rangle;$$

the standard deviation $\Delta\mathscr{A}$ is a common statistical quantity defined as a function of measurement results from an ensemble.

Often, $\Delta\mathscr{A}$ has been linked—erroneously, we believe—to the problem of compatibility by way of the Heisenberg uncertainty principle. A few remarks seem appropriate here in order to invalidate the popular contention that the uncertainty principle places restrictions on simultaneous measurability. Heisenberg's principle is a theorem, rigorously derivable from the quantum postulates; it states that under fairly general conditions,

$$\Delta\mathscr{A}\,\Delta\mathscr{B} \geq \tfrac{1}{2}|m([A, B])|,$$

where A, B are Hermitian operators representing quantum observables $\mathscr{A}$, $\mathscr{B}$, and $\Delta\mathscr{A}$, $\Delta\mathscr{B}$ refer to collectives of $\mathscr{A}$- and $\mathscr{B}$-measurements.

The principal point to be stressed here is that $\Delta\mathscr{A}$ and $\Delta\mathscr{B}$ have physical meaning only within the context of *statistics*. It is therefore illogical to interpret the uncertainty principle as a denial of the possibility of simultaneous measurement of $\mathscr{A}$ and $\mathscr{B}$ upon a single system if $[A, B] \neq 0$, as has sometimes been done. The only sense in which $\Delta\mathscr{A}\,\Delta\mathscr{B}$ may refer to a single system is purely *statistical*, i.e. to an ensemble involving *one* system sequentially measured and reprepared. Furthermore it should be noted that the product $\Delta\mathscr{A}\,\Delta\mathscr{B}$ is not even calculated from *simultaneous* measurements of $\mathscr{A}$ and $\mathscr{B}$ performed on each system. Thus, whatever conclusions one may reach 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.

To summarize: Regardless which 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 justified by its proof. Conversely, the uncertainty principle is not an a priori restriction on any consideration purely about joint measurements.

Strictly construed, the uncertainty principle is irrelevant to the problem of compatibility.

3. Trivial Joint Measurements and Commutability

There is a type of joint measurement whose consistency with quantum theory is never questioned, 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; the simultaneous measurement of a set of observables has thus been performed, albeit in a rather trivial sense. Such 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 $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$ may be used to prove that any two operators jointly measurable in this trivial sense necessarily commute.

These considerations do not imply that noncommuting observables are incompatible; they merely establish that such observables are not trivially compatible. Nevertheless, since $[A, B] = 0$ is (1) a necessary condition for trivial joint measurability of $\mathscr{A}$ and $\mathscr{B}$ and (2) the only condition under which $\Delta\mathscr{A}\ \Delta\mathscr{B} = 0$ may hold, it is sometimes claimed (via one of the misinterpretations of the uncertainty principle) that the only simultaneous measurements permitted by quantum theory are the trivial ones, that commutability is the mathematical criterion of compatibility. But in view of our preceding remarks about the uncertainty principle, such a position is evidently not tenable.

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.

Omitting the detailed proof,[16] we state here only the essential result. Let $\mathscr{A}$ and $\mathscr{B}$ be simultaneously measurable through an auxiliary variable $\mathscr{C}$ defined by

$$\mathscr{A} = \mathscr{F}(\mathscr{C}),\ \mathscr{B} = \mathscr{G}(\mathscr{C}),$$

supposing that $[A, B] = 0$.

The joint probability that $\mathscr{M}_1(\mathscr{A})$ and $\mathscr{M}_1(\mathscr{B})$ will yield (a_k, b_l) for the state ρ is then given uniquely by

$$W(a_k, b_l; \rho) = W_{\mathscr{C}}(c_{kl}; \rho) = \mathrm{Tr}\ (\rho P_{\gamma_{kl}}),$$

where the new symbols are defined through the spectral expansion of C,

$$C = \sum_{kl} c_{kl} P_{\gamma_{kl}}.$$

This analysis leads[17] to a proposition of some importance, namely: (J) The joint probability $W(a_k, b_l; \rho)$, $[A, B] = 0$, is a unique functional of the state ρ; thus the *state* of an ensemble is 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; once a W for a given ρ is found by the method of trivial joint measurement, it may be assumed that it is *the* W associated with the given ρ, independently of how $\mathscr{A}$ and $\mathscr{B}$ might be measured.

Suppose, however, that $[A, B] \neq 0$. Then the method of trivial joint measurements is of course inapplicable. Does (J) 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 later sections.

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 Sec. 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,[18] due to von Neumann, that 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[19] interested in modern physics to *define* the term "simultaneously measurable" by the commutability 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 Sec. 1. Because von Neumann's theorem is of central importance to the problem of compatibility, it is appropriate here to consider it, even though space does not permit a full analysis of the hypotheses on which it is based. It reads:

SIMULTANEOUS MEASURABILITY THEOREM

If $\mathscr{A}$ and $\mathscr{B}$ are compatible and $\mathscr{A} \leftrightarrow A$, $\mathscr{B} \leftrightarrow B$, then $[A, B] = 0$.

Expressed succinctly, the theorem says that if $\mathscr{A}$ and $\mathscr{B}$ are compatible, they are trivially compatible, for their operators necessarily commute. Unlike the semiclassical *Gedankenexperimente*, the vague interpretations of the uncertainty principle, and some strange philosophizing about subjective wave-packet reductions, the theorem offers an argument strong and clear in behalf of the proposition that noncommuting 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.

We note here merely that von Neumann's proof crucially involves the postulate of strong correspondence, discussed in Sec. 2.

Inspired by the preceding theorem, various authors[20] have suggested that quantum mechanics should be rephrased in a new *logical* framework which would properly allow for incompatibility. We believe, and intend to show in subsequent sections of this article, that von Neumann's mathematics does not in fact establish incompatibility as an intrinsic quantal property. Hence, if our analysis is correct, any "quantum logic" designed to embrace incompatibility is founded upon a mistaken interpretation of quantum physics. We shall now expose certain salient features of so-called "quantum logic," in order to establish its relation to von Neumann's theorem.

Propositions, or questions, can be introduced into quantum theory as functions of observables. Consider an observable $\mathscr{A} \leftrightarrow A = \sum_k a_k P_{\alpha_k}$ and the proposition $\mathscr{P}_n$: "$\mathscr{M}_1(\mathscr{A})$ will yield a_n." Proposition $\mathscr{P}_n$ is simply the observable "measured" as follows: Measure $\mathscr{A}$; if a_n results, assign to $\mathscr{P}_n$ the value 1; if $a_k (\neq a_n)$ emerges, assign to $\mathscr{P}_n$ the value 0. In short, $\mathscr{P}_n = \mathscr{F}_n(\mathscr{A})$, where $\mathscr{F}_n$ is defined by $\mathscr{F}_n(a_k) = \delta_{nk}$. Hence

$$\mathscr{P}_n \leftrightarrow \mathscr{F}_n(A) = \sum_k \mathscr{F}_n(a_k) P_{\alpha_k} = P_{\alpha_n}.$$

A suitable projection operator may also be found for any proposition involving commuting observables; but because of von Neumann's theorem, any compound proposition involving noncommuting observables must be regarded as undecidable, or absurd. For any two compatible propositions $\mathscr{P}$ and $\mathscr{Q}$, it is possible to find operators corresponding to the logical relation $\mathscr{P}$ "or" $\mathscr{Q} \equiv \mathscr{P} \cup \mathscr{Q}$ and $\mathscr{P}$ "and" $\mathscr{Q} \equiv \mathscr{P} \cap \mathscr{Q}$:

$$\mathscr{P} \cup \mathscr{Q} \leftrightarrow P + Q - PQ,$$

$$\mathscr{P} \cap \mathscr{Q} \leftrightarrow PQ.$$

The change in logic said to be necessitated by quantum mechanics has to do with the classical distributive law of propositions:

$$\mathscr{P} \cap (\mathscr{Q} \cup \mathscr{R}) = (\mathscr{P} \cap \mathscr{Q}) \cup (\mathscr{P} \cap \mathscr{R}).$$

Suppose, for example, that A, B are operators in a two-dimensional Hilbert space. If $[A, B] \neq 0$, and $\mathscr{P} \leftrightarrow P_{\beta_1}$, $\mathscr{Q} \leftrightarrow P_{\alpha_1}$, $\mathscr{R} \leftrightarrow P_{\alpha_2}$, then, because of von Neumann's theorem, the distributive law cannot hold in quantum theory. This stems from the correspondence

$$\mathscr{P} \cap (\mathscr{Q} \cup \mathscr{R}) \leftrightarrow P_{\beta_1}(P_{\alpha_1} + P_{\alpha_2} - P_{\alpha_1}P_{\alpha_2}) = P_{\beta_1}(1 - 0) = P_{\beta_1};$$

but $(\mathscr{P} \cap \mathscr{Q}) \cup (\mathscr{P} \cap \mathscr{R})$ is an absurd proposition, for neither $\mathscr{P} \cap \mathscr{Q}$ nor $\mathscr{P} \cap \mathscr{R}$ is measurable, since they are compounds of $\mathscr{A}$ and $\mathscr{B}$ with $[A, B] \neq 0$. Thus, since the distributive law apparently cannot hold in quantum theory, it has been suggested that some "nondistributive" logic is required for quantum propositions.

We shall return to this point in Sec. 7.

5. Counterexamples Suggesting That Noncommuting Observables Are Compatible

Mathematically, von Neumann's simultaneous measurability theorem is beyond criticism; it 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*. It is possible to construct such counterexamples, and two of them will be recorded.

Consider first the quantum theory connected with the measurement of a single observable, viz. the "time-of-flight" method for measuring the momentum $\mathscr{P}$ of an electron. The rule of correspondence for position $\mathscr{X}$ might consist, for example, of the direct observation of a coincidence between a scale mark and a macroscopic spot appearing on a photographic plate in response to an electron impact.

An "electron gun" prepares the state $\rho = P_\psi$. Using non-relativistic wave mechanics, we find that the probability density $w_{\mathscr{P}}(p; \psi)$ for $\mathscr{M}_1(\mathscr{P})$ at the time of preparation is

$$w_{\mathscr{P}}(p; \psi) = \left| \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \psi(x)\, dx \right|^2 .$$

This distribution is the quantum-mechanical test for deciding whether a proposed experiment which generates numbers via the established operational definition for $\mathscr{M}_1(\mathscr{X})$ qualifies as a momentum measurement scheme $\mathscr{M}_2(\mathscr{P})$. If the numbers in question are to be regarded as $\mathscr{M}_1(\mathscr{P})$-results, they must satisfy the theoretical distribution $w_{\mathscr{P}}(p; \psi)$.

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 $\psi(x) \not\equiv 0$ is $(-x_0, x_0)$. The $\mathscr{M}_2(\mathscr{P})$-procedure[21] is simple: We simply wait a very long time $(t \to \infty)$ as the electron moves *freely*, and then we measure the observable $\mathscr{F}(\mathscr{X}) = m\mathscr{X}/t$, where m is the electron mass. The number obtained then counts as the result of $\mathscr{M}_1(\mathscr{P})$ at $t = 0$. To justify this operational definition of $\mathscr{P}$ quantum mechanically, one must prove that the probability for $\mathscr{M}_1(\mathscr{P})$ to yield $p \in (p_1, p_2)$ at $t = 0$ equals the probability that $\mathscr{M}_1[\mathscr{F}(\mathscr{X})]$ yields $(m\mathscr{X}/t) \in (p_1, p_2)$ at $t \to \infty$. The details are given in Ref. 9.

We conclude that the results of "direct" $\mathscr{F}(\mathscr{X})$-measurements, performed sufficiently long after the preparation of $\psi(x, 0)$, will be distributed just like the theoretical results for $\mathscr{M}_1(\mathscr{P})$ upon $\psi(x, 0)$. This time-of-flight arrangement is therefore fully certified quantum mechanically as an operational definition of $\mathscr{P}$. Because quantum theory can make only statistical predictions, no further guarantee that this method "really" makes $\mathscr{P}$-measurements is required. Indeed, further *quantal* analysis of the question is theoretically meaningless.

One can also see that this time-of-flight method for obtaining the results which $\mathscr{M}_1(\mathscr{P})$ at $t = 0$ would yield, determines likewise the result which $\mathscr{M}_1(\mathscr{P})$ would yield any time $t > 0$. This follows at once from the fact that momentum is conserved in the free motion of the electron; in quantum-mechanical terms,

$$W_{\mathscr{P}}[p \in (p_1, p_2); \psi(x, 0)] = W_{\mathscr{P}}[p \in (p_1, p_2); \psi(x, t)].$$

Hence, by the same reasoning which validated the time-of-flight method as a rule of correspondence for $\mathscr{M}_1(\mathscr{P})$ at $t = 0$, we can regard the results of $\mathscr{M}_1[\mathscr{F}(\mathscr{X})]$, $t \to \infty$, as $\mathscr{M}_1(\mathscr{P})$-results for *any* $t > 0$. In particular, consider the instant when the electron strikes the photographic plate and the result emerges. For that instant we may conclude with full quantum-mechanical justification that $\mathscr{M}_1(\mathscr{P})$ would have yielded $\mathscr{F}(x)$ where x is the result of the $\mathscr{M}_1(\mathscr{X})$. Contrary to the prohibitions of von Neumann's theorem, we have here an empirical method for the simultaneous measurement of $\mathscr{X}$ *and* $\mathscr{P}$, two renowned noncommuting observables!

There is a tendency to dismiss simultaneous measurement schemes such as the one just described as if they did not in fact *legitimately* challenge the orthodox view. One authoritative argument was first employed by Heisenberg and may be summarized by his statement[22] that "the uncertainty relation does not refer to the past." In the time-of-flight experiment, by the time the $\mathscr{X}$, $\mathscr{P}$-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, 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."[23]

In rejoinder to this distinctly philosophical and somewhat subjective disposal of the matter, we offer the following comments.

The word *knowledge* is not unambiguous when employed in discussions regarding quantum measurement. As we have seen, from a strict quantal point of view, an electron *never possesses properties* $\mathscr{X}$, $\mathscr{P}$ of which one can be knowledgeable or ignorant. (There does not exist a preparation scheme Π which produces electrons always yielding the same measured $\mathscr{X}$, $\mathscr{P}$-values.) Accordingly, measurement should never be described as though it increased knowledge by revealing perhaps with growing precision the actual, previously unknown, "value" of an observable. Measurements merely generate numerical results associated with certain operations upon the system of interest. The meaning of these numbers is provided by the theory into which they are fed; in quantum theory the numbers are not to be regarded as measures of possessed attributes.

It is therefore not very meaningful 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 and have no bearing on measurements upon a single system at a single time, since standard deviations refer only to measurements upon ensembles. Hence, as already explained in Sec. 2, the emergence of simultaneous $\mathscr{X}$, $\mathscr{P}$-values upon measurement in no way violates the uncertainty principle.

In the time-of-flight method, the $\mathscr{X}$, $\mathscr{P}$-measurement results-admittedly refer to the instant just prior to the electron's impact on the plate. They are indeed useless for predicting (in classical fashion) the result of a future $\mathscr{X}$-measurement, yet 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. Reference of the $\mathscr{X}$, $\mathscr{P}$-values to a past time is no special feature of *simultaneous* measurements; it is characteristic of all quantum measurements. Certainly, the time-of-flight measurement of $\mathscr{P}$ alone referred to $t = 0$, although the result did not emerge until $t \to \infty$. Nevertheless, such $\mathscr{P}$-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 $\psi(x, 0)$. Surely, if the physical significance of such $\mathscr{X}$, $\mathscr{P}$-values is a matter of "personal belief," then all measurement results for *single* observables are likewise merely of solipsistic significance.

We are therefore forced to conclude that the foregoing method for simultaneous measurement of $\mathscr{X}$, $\mathscr{P}$ is as significant as any other quantum-mechanical measurement scheme.

We now present another counterexample to the simultaneous measurability theorem. Consider two quantum systems S_1, S_2 with observables $\mathscr{A}_1$, $\mathscr{B}_1$, and $\mathscr{A}_2$ associated with S_1 and S_2, respectively. Let $[A_1, B_1] \neq 0$ and denote eigenvectors and eigenvalues as follows:

$$A_1\alpha_k^{(1)} = a_k^{(1)}\alpha_k^{(1)}, \qquad A_2\alpha_k^{(2)} = a_k^{(2)}\alpha_k^{(2)}.$$

Although S_1 and S_2 are noninteracting, they are assumed to be in a correlated state:

$$\Psi = \sum_k c_k\alpha_k^{(1)} \otimes \alpha_k^{(2)}.$$

If $\mathscr{A}_2$ has an operational definition, the correlation in Ψ that relates $\mathscr{M}_1(\mathscr{A}_1)$-results to $\mathscr{M}_1(\mathscr{A}_2)$-results may be used to construct an $\mathscr{M}_2(\mathscr{A}_1)$. As in the $\mathscr{M}_2(\mathscr{P})$ case of our previous example, we must establish a theoretical matching between probabilities associated with $\mathscr{M}_1(\mathscr{A}_2)$ and $\mathscr{M}_1(\mathscr{A}_1)$. Since

$$[A_1, A_2] = [A_1 \otimes 1, 1 \otimes A_2] = 0,$$

$\mathscr{A}_1$ and $\mathscr{A}_2$ may be jointly measured (trivially) through an auxiliary observable (cf. Sec. 3). The joint probability $W(a_k^{(1)}, a_l^{(2)}; \Psi)$ is then

$$W(a_k^{(1)}, a_l^{(2)}; \Psi) = \operatorname{Tr}(P_\Psi P_{\alpha_k^{(1)} \otimes \alpha_l^{(2)}}) = |c_k|^2\, \delta_{kl}.$$

From this expression it is apparent that when $\mathscr{M}_1(\mathscr{A}_2)$ yields $a_k^{(2)}$, a simultaneous $\mathscr{M}_1(\mathscr{A}_1)$ would yield $a_k^{(1)}$. Hence we have an

$\mathscr{M}_2(\mathscr{A}_1)$ scheme: To measure $\mathscr{A}_1$, simply measure $\mathscr{A}_2$; if $a_k^{(2)}$ results, then $a_k^{(1)}$ is regarded as the result of $\mathscr{M}_1(\mathscr{A}_1)$.

Suppose $\mathscr{B}_1$, like $\mathscr{A}_2$, also has an established operational definition. Now, since the $\mathscr{M}_2(\mathscr{A}_1)$ just outlined involves no interaction with S_1, we may perform $\mathscr{M}_1(\mathscr{B}_1)$ simultaneously with $\mathscr{M}_2(\mathscr{A}_1)$ and thereby jointly measure noncommuting observables $\mathscr{A}_1$ and $\mathscr{B}_1$. Once again von Neumann's theorem is contradicted.

6. Strong Correspondence—the Root of Quantum Inconsistencies

Three conclusions may be drawn from the last two sections. (1) The standard quantum postulates (P1S, etc.) rigorously imply that noncommuting observables are incompatible. (2) The same postulates permit empirical arrangements which must be regarded as legitimate schemes for the simultaneous measurement of noncommuting observables, provided the term "measurement" is used in its normal sense. (3) *Hence the standard postulates of quantum theory are inconsistent.* We must therefore reexamine the axiomatic basis of von Neumann's simultaneous measurability theorem in order to discover the false hypothesis which enables the rigorous deduction of this false theorem.

Any theory about the simultaneous measurement of several observables, from axioms referring only to measurements of single observables $\mathscr{M}_1(\mathscr{A})$, requires the notion of compound observable (Secs. 2 and 4); and this concept is subject to consistency conditions (C_1) and (C_2), which would have to be satisfied by any operator corresponding to such a compound observable.

Let us examine the correspondence $\mathscr{A} + \mathscr{B} \leftrightarrow A + B$. Condition ($C_1$) alone *implies* this rule.[24] Explicitly, P1S guarantees the existence of an operator corresponding to the observable $\mathscr{A} + \mathscr{B}$; that operator would satisfy (C_1) and (C_2). But (C_1) for $\mathscr{A} + \mathscr{B}$ can be satisfied by only *one* operator, namely $A + B$. Condition (C_2), therefore, *need not be used at all.* This observation provides an important clue in our search for the false hypothesis in question.

We ask: Is $\mathscr{A} + \mathscr{B}$ truly an observable? If not, P1S cannot be invoked to assure the existence of an operator counterpart. To answer the question, we recall the last example of Sec. 5. It showed how two systems S_1 and S_2 in a realizable state could be used to construct an appropriate rule of correspondence for simultaneous $\mathscr{M}_1(\mathscr{A}_1)$ and $\mathscr{M}_1(\mathscr{B}_1)$. Now the experimenter is obviously free to add the two results; hence it is apparent that $\mathscr{A}_1 + \mathscr{B}_1$ is observable. *Therefore*, if P1S is true, *there must exist an operator S such that $\mathscr{A} + \mathscr{B} \leftrightarrow S$* in general.

The following simple, contravening example defeats this claim. Let system S_1 be a "spin" whose relevant states and operators span a two-dimensional spinor space. For our noncommuting observables $\mathscr{A}_1$, $\mathscr{B}_1$, we take x- and z-components of spin, $\mathscr{S}_x$ and $\mathscr{S}_z$. Thus, in the Pauli representation,

$$\mathscr{A}_1 = \mathscr{S}_x \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathscr{B}_1 = \mathscr{S}_z \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Now simultaneous measurements of $\mathscr{A}_1$ and $\mathscr{B}_1$ employing the correlation with the auxiliary system S_2 will, by Th. 4, always yield one of the eigenvalue pairs: $(\hbar/2, \hbar/2)$, $(\hbar/2, -\hbar/2)$, $(-\hbar/2, \hbar/2)$, $(-\hbar/2, -\hbar/2)$. Hence only the three values $\hbar, 0, -\hbar$ are possible for results of measuring $\mathscr{A}_1 + \mathscr{B}_1$. To use the set notation of Sec. 2,

$$\mathscr{N}(\mathscr{A}_1 + \mathscr{B}_1) = \{-\hbar, 0, \hbar\};$$

and by condition (C_2), if $\mathscr{A}_1 + \mathscr{B}_1 \leftrightarrow S$, quantum mechanics would be self-contradictory unless

$$(C_2) \quad \mathscr{E}(S) \subseteq \mathscr{N}(\mathscr{A}_1 + \mathscr{B}_1).$$

However, (C_1) must also be satisfied by S. As shown in Sec. 4, the only S meeting this requirement is

$$S = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

But an elementary calculation reveals that the eigenvalues of this operator are $\hbar/2^{1/2}$, $-\hbar/2^{1/2}$; in other words, the set $\mathscr{E}(S) =$

$\{-\hbar/2^{1/2}, \hbar/2^{1/2}\}$. Comparing $\mathcal{N}(\mathcal{A}_1 + \mathcal{B}_1)$ and $\mathcal{E}(S)$, with $\mathcal{A}_1 + \mathcal{B}_1 \leftrightarrow S$, one finds that

$$\mathcal{N}(\mathcal{A}_1 + \mathcal{B}_1) \cap \mathcal{E}(S) = \varnothing.$$

We conclude that the only operator S capable of satisfying (C_1) does not satisfy (C_2).

To summarize: $\mathcal{A}_1 + \mathcal{B}_1$ is demonstrably observable. Axiom P1S then ensures the existence of $S \leftrightarrow \mathcal{A}_1 + \mathcal{B}_1$. If the quantal axioms are consistent, that S must satisfy both (C_1) and (C_2). The *unique* S which satisfies (C_1) violates (C_2). Hence the axioms P1S and P2 are inconsistent.

This conclusion was already suggested at the beginning of this section upon confrontation of von Neumann's theorem with the counterexamples of Sec. 5. But we now see where the difficulty lies among the initial hypotheses leading to that theorem. The theorem is false because P1S—*strong correspondence*—proclaims the existence of operator-observable correspondences which cannot exist in harmony with the remaining postulates. Thus the axiom set—P1S, P2—must be altered. The modification we propose is to replace P1S by P1. Further corroborative evidence for the need of this change (e.g. Temple's theorem) is presented elsewhere.[25]

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 as it will not cover all empirical procedures. The historian may console himself in thinking that a fuller range of applicability may be provided by future theoretical discoveries.

At the moment we ask the nonspeculative question: What effect does the replacement of strong by weak correspondence have on the principal quantum theorems? Clearly Th. 1, for example, which implies that every real linear functional $m(A)$ of the Hermitian operators may be expressed in the form Tr (ρA), 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 *are* involved do represent observables. Within the mathematical framework, operational definitions are irrelevant, and quantum mechanics is a set of mathematical objects subject to given rules.* Their application to the world is made possible by the discovery of rules of correspondence with the *P*-field of experience.[26] None of the parts of linear algebra which form the foundation of quantum theory will be affected by the elimination of strong correspondence. In fact, a careful search through quantum theory for a proposition dependent upon strong correspondence convinced the present writers that no basic theorem involving the analysis of ensembles, statistics of measurement results, etc., requires P1S rather than P1 in its proof.

If only weak correspondence is adopted, the physicist cannot demand in a priori fashion that the mathematician furnish an F for every one of his $\mathscr{F}$'s. Given an observable $\mathscr{F}$, the operator algebra is not *expected* to produce an F; instead, it is simply *asked* whether or not F does exist such that $\mathscr{F} \leftrightarrow F$. In short, what were formerly regarded as "correspondence theorems" are now interpreted as tests of validity for proposed correspondences.

We now offer a summary of the correct interpretation of the theorems of this kind which were mentioned in previous sections:

(1) $\mathscr{A} + \mathscr{B} \leftrightarrow A + B$: (C_1) uniquely determines the operator $A + B$ but (C_2) is often violated. The correspondence is therefore not generally valid.

(2) $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$: Von Neumann's theorem (Sec. 4), in the proof of which this correspondence was central, actually demonstrates that this correspondence can apply only to commuting operators (in which case it takes the simple form $\mathscr{A}\mathscr{B} \leftrightarrow AB = BA$).

(3) What corresponds to $\mathscr{A}\mathscr{B}\mathscr{C}$? Temple's theorem exhibits the ambiguities inherent in this triple product when P1S is assumed.[27]

* Among these are tacit rules concerning the construct $\mathscr{M}_1(\mathscr{A})$ which give meaning to the primitive term observable as it appears in P1.

It is perhaps instructive to consider a simple example which illustrates why consistency condition (C_2) required only $\mathscr{E}(F) \subseteq \mathscr{N}[\mathscr{F}(\mathscr{A}, \mathscr{B})]$ and not the equality of these two sets. Suppose $\mathscr{A} = \mathscr{L}_z^2$, $\mathscr{B} = \hbar\mathscr{L}_z$, where $\mathscr{L}_z$ is the z-component of orbital angular momentum, $\mathscr{L}_z \leftrightarrow L_z = (\hbar/i)\,\partial/\partial\varphi$. $\mathscr{A}$ and $\mathscr{B}$ are measured simultaneously, and the results are added together. The set of all possible results of this procedure, $\mathscr{N}(\mathscr{A} + \mathscr{B})$, is given by $\mathscr{N}(\mathscr{L}_z^2 + \hbar\mathscr{L}_z) = \{m^2\hbar^2 + n\hbar^2\}$, since $\mathscr{E}(L_z) = \{m\hbar\}$. But the eigenvalues of $L_z^2 + \hbar L_z$ comprise the set $\mathscr{E}(L_z^2 + \hbar L_z) = \{k(k + 1)\hbar^2\}$, which is only a subset of $\mathscr{N}(\mathscr{A} + \mathscr{B})$. The reason for this inequality is easily understood if postulate (J) of Sec. 3 is recalled. Any measurement of the observables $\mathscr{L}_z^2$ and $\hbar\mathscr{L}_z$ must yield results correlated in the same manner as would be the results of a trivial joint measurement of these observables. One such joint measurement involves simply measuring $\mathscr{L}_z$ and evaluating $\mathscr{L}_z^2 + \hbar\mathscr{L}_z$. But that procedure can yield only numbers in the set $\{k(k + 1)\hbar^2\} = \mathscr{E}(L_z^2 + \hbar L_z)$. This demonstration merely affirms the consistency of (J), with the postulated correspondence $\mathscr{F}(\mathscr{C}) \leftrightarrow \mathscr{F}(C)$.

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 false for most $\mathscr{A}$ and $\mathscr{B}$, it is evident that so-called "quantization" schemes based upon (1) and (2) are at best memory aids taking advantage of our familiarity with classical mechanics. (Cf. Ref. 9.)

While replacement of P1S by P1 has no effect whatsoever on the normal applications of the theory to experiment, this revision does have considerable theoretical and philosophical significance. Primarily it shows that von Neumann's simultaneous measurability theorem is a correct mathematical theorem physically misinterpreted as a restriction on measurability. It turns out to be a *reductio ad absurdum* proof that the correspondence $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$ is false unless $[A, B] = 0$; in other words, a proof that $[A, B] = 0$ is a necessary condition for the validity of $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$.

Hence any physical or metaphysical idea motivated by, or founded upon, the concept of incompatibility now requires careful reexamination. Two very common propositions based on incompatibility 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 outcome of simultaneous measurements of noncommuting observables is meaningless, a new system of logic is required for quantum physics.

(1) When the incompatibility doctrine has been discarded, there remains no a priori restraint against the study of joint distributions. (For a systematic study of such distributions, cf. Ref. 9.)

(2) At the end of Sec. 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

$$\mathscr{P} \cap (\mathscr{Q} \cup \mathscr{R}) = (\mathscr{P} \cap \mathscr{Q}) \cup (\mathscr{P} \cap \mathscr{R}),$$

which merely expresses an idea most physicists—including quantum theorists "off duty," to use Landé's phrase—regard as "common sense." The problem was that propositions $\mathscr{P}$, $\mathscr{Q}$, and $\mathscr{R}$ can be given for which there does exist an Hermitian 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 $\mathscr{A} + \mathscr{B} \leftrightarrow S$. Just as an appropriate S exists only when $[A, B] = 0$, similarly a D exists such that $\mathscr{P} \cap \mathscr{Q} \leftrightarrow D$ only when $[P, Q] = 0$. When $[P, Q] \neq 0$, it simply means that the compound proposition $\mathscr{P} \cap \mathscr{Q}$ 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!

There are other interesting implications with respect to the "microcausality principle." They are discussed in Ref. 9.

8. A Search for "Simple" Simultaneous Measurements

It is shown elsewhere [28–30] that attempts to approach the study of quantum joint probabilities of noncommuting observables via more or less natural random-variable techniques are thwarted at some stage by ignorance of, or perhaps even the nonexistence of, operators corresponding to compound observables. It is therefore desirable 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 Sec. 1. It was seen there that the primitive classical notion of possession ("System S has $\mathscr{A}$-value a_k") is superseded by the primitive quantal measurement construct $\mathscr{M}_1$ ("If $\mathscr{M}_1(\mathscr{A})$ is performed on system S, the value a_k will result with probability . . ."). While 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 $\mathscr{M}_1$'s. On the other hand, *statements* of such connections and associated empirical procedures constitute the usual scientific concept of measurement, or measurement scheme (operational definition, the epistemic correspondence rule[31]). To signalize the logical distinction, we have designated the latter class of constructs, which form part of the *theoretical* structure of our problem, by $\mathscr{M}_2$. They were exemplified in Sec. 5.

Thus far we have shown by way of examples that there *are* procedures which permit an assignment of values to pairs of noncommuting observables. Our present aim goes beyond such indications; it is to clarify within the context of measurement *theory*, as presented in the foregoing pages, how such empirical operations function as parts of the complete mathematical structures. We shall see that certain kinds of $\mathscr{M}_2$ are free from theoretical difficulties, while others seem to generate internal contradictions.

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 *derivable* within the framework of a quantum theory of $\mathcal{M}_2$. That is, if a given procedure $\mathcal{M}_2(\mathcal{X}, \mathcal{P})$ is to be regarded as a method for simultaneous measurement of $\mathcal{X}$ and $\mathcal{P}$, the scheme must be certified by a theory establishing relations between $\mathcal{M}_1(\mathcal{X})$, $\mathcal{M}_1(\mathcal{P})$ and whatever "direct meter readings" are used as the basis for inference of simultaneous $\mathcal{M}_1(\mathcal{X})$- and $\mathcal{M}_1(\mathcal{P})$-results; from this analysis it should be possible in principle to find the probability for the occurrence of those "meter readings" that imply any given pair of $\mathcal{X}$- and $\mathcal{P}$-values. This measurement-theoretical approach to the joint-probability problem bypasses the difficulty associated with the operator-observable correspondence, 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 $\mathcal{M}_2$-concepts: (1) simple or type A and (2) historical or type B. This distinction will later turn out to have considerable bearing on the problem of compatibility.

(1) A simple $\mathcal{M}_2$ begins with system S in an arbitrary state ρ_{t_0} at some specified time t_0 and demonstrates how some single operation upon S eventually leads to numbers from which may be inferred $\mathcal{M}_1$-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. We shall also refer to this class of measurement as belonging to type A.

(2) An *historical* $\mathcal{M}_2$-theory also seeks to certify some operation as a bona fide supplier of numbers which can be meaningfully interpreted as $\mathcal{M}_1$-results for S in state ρ_{t_0}. However, unlike the simple type A, the historical $\mathcal{M}_2$-theory cannot be worked out without detailed information concerning the structure of ρ_{t_0}. Such information might be deduced from facts about the 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 Sec. 5: the simple time-of-flight $\mathcal{M}_2(\mathcal{P})$ of type A and the historical time-of-flight $\mathcal{M}_2(\mathcal{X}, \mathcal{P})$ of type B.

Physically, the $\mathcal{M}_2$-theories of type A have been of greatest interest, because they represent the idea of measurement in its most primitive form, as a process applicable to a system at any instant independently of its past. Similarly, in quantum mechanics the language of $\mathcal{M}_1$'s tends to presuppose that measurements are performed upon systems in states which are simply given without details as to the actual method of preparation. Accordingly, $\mathcal{M}_2$-schemes for single observables (or commuting sets of observables) have been of type A. One might therefore be tempted to seek a simple $\mathcal{M}_2$-theory covering the simultaneous measurement of several noncommuting observables. However, in view of the fact that both examples of simultaneous measurement given in Sec. 5—the time-of-flight $\mathcal{M}_2(\mathcal{X}, \mathcal{P})$ and the use of two systems already correlated at the time of interest—were of type B, there is so far no reason to expect any *simple* theory for simultaneous measurement.

Elsewhere[32] we have examined two fairly general procedures which, at the outset, seem to be altogether plausible methods for achieving simultaneous type A measurement of two noncommuting observables. In both cases, theoretical obstacles eventually arose, and this may be interpreted as evidence that quantum theory does perhaps forbid *type A* simultaneous measurements. Deeper reasons to anticipate such a theoretical prohibition have also been explored.[33]

These examples furnish partial evidence for this proposition:

(0)
Simultaneous type A measurements of noncommuting observables are theoretically impossible.

Of course, merely citing two unsuccessful attempts to develop a simple $\mathcal{M}_2(\mathcal{A}, \mathcal{B})$ does not prove (0); nevertheless there appears, for the first time in the present study, good reason to suspect that quantum theory may indeed place some restriction upon joint measurability. If so, the qualification will not be a sweeping mandate to the effect that $\mathcal{M}_2(\mathcal{A}, \mathcal{B})$ is generally impossible, since that common version was refuted in Sec. 5 by counterexamples. Rather,

(0) would mean only this: Given at time t_0 a system S *of unknown history*, it is impossible to devise an operation $\mathcal{M}_2(\mathcal{A}, \mathcal{B})$ which leads to numbers (a_k, b_l) interpretable as $\mathcal{M}_1(\mathcal{A})$- and $\mathcal{M}_1(\mathcal{B})$-results for time t_0.

References

1.
A. Landé, *New Foundations of Quantum Mechanics* (Cambridge University Press, 1965), p. 124.
2.
J. L. Park, Am. J. Phys. **36**, 211 (1968).
3.
J. L. Park, Phil. Sci. **35**, pt. I, 205; pt. II, 389 (1968).
4.
H. Margenau, Phil. Sci. **30**, 6 (1963).
5.
H. Margenau, Phys. Rev. **49**, 240 (1936).
6.
J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, trans. by R. T. Beyer (Princeton University Press, 1955).
7.
Ref. 4.
8.
Ref. 3.
9.
J. L. Park and H. Margenau, Intern. J. Theoret. Phys. **1**, 211 (1968). This paper contains a fuller account and more detailed proofs of many of the results set forth in the present article.
10.
H. Margenau, Phys. Today **7**, 6 (1954).
11.
H. Margenau, Phil. Sci. **4**, 352–356 (1937).
12.
G. C. Wick, A. S. Wightman, and E. P. Wigner, Phys. Rev. **88**, 101 (1952).
13.
Ref. 6, pp. 313–316. Cf. also A. M. Gleason, J. Math. Mech. **6**, 885 (1957).
14.
See, for instance, H. J. Groenewold, Physica **12**, 405 (1946); J. R. Shewell, Am. J. Phys. **27**, 16 (1959).
15.
Ref. 9.
16.
J. von Neumann, Ann. Math. **32**, 191 (1931).

17.
Ref. 9.
18.
Ref. 6, pp. 225–230.
19.
E.g.: G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, 1963), p. 70; I. E. Segal, Ann. Math. **48**, 930–948 (1947); V. S. Varadarajan, Commun. Pure Appl. Math. **15**, 189–217 (1962).
20.
E.g.: G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936); H. Reichenbach, *Philosophic Foundations of Quantum Mechanics* (University of California Press, 1944); C. Piron, Helv. Phys. Acta **37**, 439 (1964).
21.
R. P. Feynman, *Quantum Mechanics and Path Integrals* (McGraw-Hill, 1965), pp. 96–98.
22.
W. Heisenberg, *The Physical Principles of the Quantum Theory* (University of Chicago Press, 1930), p. 20.
23.
Ibid.
24.
Ref. 9.
25.
Ibid.
26.
Cf. H. Margenau, *The Nature of Physical Reality* (McGraw-Hill, 1950), pp. 171–177.
27.
Ref. 9.
28.
H. Margenau and R. N. Hill, Progr. Theoret. Phys. (Kyoto) **26**, 727 (1961).
29.
H. Margenau, Ann. Phys. (N.Y.) **23**, 469 (1963).
30.
L. Cohen, J. Math. Phys. **7**, 781 (1966); H. Margenau and L. Cohen in *Quantum Theory and Reality*, ed. by M. Bunge (Springer-Verlag, 1968).
31.
H. Margenau, Phil. Sci. **2**, 1 (1935).
32.
Ref. 9.
33.
Ibid.