International Journal of Theoretical Physics, Vol. 1, No. 3 (1968), pp. 211–283.

Simultaneous Measurability in Quantum Theory†

JAMES L. PARK‡ and HENRY MARGENAU
Yale University, New Haven, Connecticut

Abstract§

This paper presents a study of what is sometimes regarded as the conceptual heart of quantum theory, namely, 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 elsewhere (Park, 1967b). Contrary to custom in discussions on simultaneous measurability, the uncertainty principle is not dwelt upon but simply stated carefully in order 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 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

† Part of a dissertation presented by James L. Park to the faculty of the Graduate School of Yale University in candidacy for the degree of Doctor of Philosophy. Research supported in part by the Air Force Office of Scientific Research under Grant 843-67.

‡ Present address: Washington State University, Pullman, Washington.

§ Script letters ($\mathscr{A}, \mathscr{B}$) are used to designate observables; capital italic letters (A, B) denote operators; lower-case italic letters (a, b) denote numerical measured values; and lower-case Greek letters (α, β), except ρ, denote quantum state vectors (ρ is used for the density operator).

(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, namely, 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 on 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 (1949) and Heisenberg (1930), 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 (Jamme, 1966).

(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 if 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 $\delta l = k/\delta w$, 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 (naïve 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 (Mandl, 1957; Roman, 1965). 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

† 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 [Margenau, 1963; Park, 1967b; and particularly *Physical Review*, **49**, 240 (1936)]. This point is further elaborated by Park (1968a, b). Briefly stated, our premise here is that every measurement determines a number which characterizes *one* element of an ensemble. *Many* such numbers define the distribution of probabilities over the ensemble. Since a state vector is the equivalent of a *set* of distributions, a single measurement can not in general determine a state. If it did, quantum mechanics would be a freak when viewed as an example of a stochastic theory. The projection postulate assumes such a miracle. There are single measurements which, when coupled with extraneous information which does not arise analytically from the axioms of quantum mechanics, may allow the establishment of quantum states; but these are special cases. (For example if a system is known to have been isolated for a long time and I measure its energy only once, its state is fixed. This is like knowing that an institution has inmates of the same sex. A single 'measurement' of the sex of an inmate then settles the distribution.)

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* (Park, 1968a) 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 (Margenau, 1937; 1963b; Park, 1968b).

(4) *Projection Postulate (von Neumann's measurement intervention transformation)* (von Neumann, 1955). There is a way (Margenau, 1963b) 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 (Park, 1968b). 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† (Section 4).

As stated above, the purpose of the present work is to scrutinize and evaluate the principle of incompatibility of noncommuting 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

† Of course the projection postulate does appear in von Neumann's book, but it plays no role in the theorem here considered.

be mentioned again. In particular, no further discussion will appear concerning (1), (3), and (4). The value of semiclassical gedanken-experiments (1) has already been commented upon; and the absurdity of the naïve projection postulate (3) and the inadequacy, or lack of universality, of its sometimes correct version (4) are fully discussed elsewhere by one of the present authors (Park, 1968a, b). 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. This peculiarity of quantum observables has been characterized by one of us as *latency* (Margenau, 1950; 1954). A brief digression in explanation of the idea of latency seems appropriate at this point.

Traditional philosophy features a distinction between primary and secondary qualities, an issue which recalls such figures as Galileo and Locke. It relates to a division of observables into two categories, those which, like size, shape and position, adhere intrinsically to physical objects and those which arise in the process of perception, e.g. color, taste, and smell. This issue may now be regarded as dead because the latter class has been recognized as the subjective ('protocol') counterpart of the former. There are, to be sure, primary qualities (like magnetic fields) to which, because of our limited sensory equipment, no secondary qualities correspond, hence there is no unique correspondence between primary and secondary qualities. Nevertheless a distinction between observables (qualities) which adhere to the objects of inquiry and others which do not has been revived by the quantum theory. The latter, however, are no longer secondary but latent. They do not arise in the act of observation but become concrete or actualized or numerical when the system is subjected to measurement.

A free electron in an energy eigenstate yields a definite number when its position is measured. Yet a similar measurement on it when the same state has been reprepared may yield another value: There is

a dispersion of measured values. As we have noted, the classical tendency is to account for this dispersion by invoking the disturbance of the state of the electron by the measuring device. This, for many reasons (some of which have already been cited), is an erroneous explanation which attempts the impossible logical feat of deriving the more general theory (quantum mechanics) from its own special case (classical mechanics). Certainly, disturbance by measuring devices can be used to exemplify the quantal dispersion, but not to explain it in its universality. The origin of the dispersions lies deeper than that. It is not proper to suppose that the electron in a state of the Schrödinger form $\exp(ipx)$ *has* a position which we merely do not know until we measure it. Its ability to interfere with itself clearly implies that 'it is in many places at once'. In other words, it does not *have* a position, position is not a *possessed* observable. We call it a *latent* one. A latent observable does not in general have a value at all; it assumes one when a measurement elicits it. It is clear, of course, that an observable which is generally latent can under special circumstances be regarded as possessed, namely when the system is in an eigenstate of that observable. We feel that the acknowledgment of the existence of latent qualities in nature is far safer, simpler and less compromising than the usual talk about complementarity, which becomes unnecessary when latency as a primitive idea is accepted.

The change, brought about by the quantum theory, then, is this. The old primary qualities correspond to possessed observables, those which, when they reveal themselves in repeated measurements on systems identically prepared or, when possible, on the same systems in their continuing undisturbed states, do not scatter their values. Secondary qualities are replaced by latent ones which, in a certain obscure but primitive sense, 'come about' or 'spring into being' when a measurement is made, yet allow no inference with respect to numerical values present in the premeasurement state. At the present time elementary quantum theory recognizes a few possessed observables such as the mass, the charge and the (absolute value of the) spin of an electron; its position, momentum, energy, vector spin and most others are latent. In classical philosophy, every secondary quality had a primary counterpart. Latent observables, so far as is now known, do not point to possessed companions.

In a speculative and unrigorous manner of arguing one might say that the discovery of latent qualities in quantum mechanics is merely the recognition of a situation that is commonplace in the behavioral sciences. Happiness, equanimity, anger are observable qualities of man, but they are latent, not necessarily present or numerically

descriptive of a person at all times. And they, too, can spring into actuality by a psychological measurement, a certain felicitous or infelicitous act of inquiry. But we do not wish to press this analogy too far.

A more proper question which arises here is this. Are not perhaps all physical observables in the strictest sense and in their last analysis latent ones? Clearly if there is a spectrum of masses, charges and spins, and thus for each an operator with eigenvalues equal to the values of these observables now known, they too become latent. We tend to believe that a future theory will show all observables to be ultimately latent ones, and that the concept of a possessed quality is nothing but a limiting case of the former.

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 (Margenau, 1937).

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 $\mathscr{A}$ 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 $\mathscr{M}_1(\mathscr{A})$.

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 $\mathscr{M}_1$. 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 $\mathscr{M}_2$. In short, $\mathscr{M}_2$ is the normal measurement construct of all science; $\mathscr{M}_2(\mathscr{A},\mathscr{B},\ldots)$ represents any empirical procedure yielding numbers $a, b, \ldots$ which through a theory can be interpreted as the values associated with observables $\mathscr{A}, \mathscr{B}, \ldots$. In quantum theory, the values are associated with the observables via $\mathscr{M}_1(\mathscr{A})$, $\mathscr{M}_1(\mathscr{B})\ldots$ 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 $\mathscr{M}_1$ (needle position) would yield certain values and the probabilities that $\mathscr{M}_1$ (speed of system) would yield certain values. A measurement scheme $\mathscr{M}_2$ is thus explained quantum theoretically *in terms of* the primitive measurement construct $\mathscr{M}_1$.

Using these concepts, it is possible to define precisely what is meant by simultaneous measurability of two observables: *observables $\mathscr{A}$ and $\mathscr{B}$ will be termed compatible, simultaneously measurable, or jointly measurable if there exists an $\mathscr{M}_2(\mathscr{A},\mathscr{B})$*, i.e., an operation yielding two numbers a, b with the same probabilities that quantum theory confers upon the two propositions '$\mathscr{M}_1(\mathscr{A})$ yields a' and '$\mathscr{M}_1(\mathscr{B})$ yields b', where both $\mathscr{M}_1$'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 $\mathscr{A},\mathscr{B}$ are noncommuting observables, is it quantum theoretically possible for an $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ 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 $\mathscr{A}$, then the operator $\mathscr{F}(A)$ corresponds to observable $\mathscr{F}(\mathscr{A})$, where $\mathscr{F}$ is a function.

It is convenient to use the symbol $\leftrightarrow$ to represent this operator-observable correspondence relation; thus $A \leftrightarrow \mathscr{A}$ means A 'corresponds to' $\mathscr{A}$ in the sense of P1. The observable $\mathscr{F}(\mathscr{A})$ is defined by this measurement procedure: 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

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

(Extension to degenerate and/or continuous spectra is straightforward.)

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 $A \leftrightarrow \mathscr{A}$, 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 that of their analogues in typical von Neumann-inspired axiomatizations (von Neumann, 1955). 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. Wick, Wightman and Wigner (1952) 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

† In postulates and theorems, the term '$\mathscr{A}$-measurement' refers to the primary quantum measurement construct $\mathscr{M}_1(\mathscr{A})$.

‡ In effect, these superselection rules keep certain subspaces of the full Hilbert space apart by fiat, in the sense that there can never be a physical state which is a superposition of vectors from different subspaces. For instance, no angular momentum state can be a superposition of integral and half integral J. (In this instance, the incompatibility of the spaces can be proved on the basis of relativistic invariance.) The relevance of superselection rules from the present point of view is this. It is mathematically possible to construct the forbidden superposition and presumably to formulate an Hermitean operator for which it is an eigenstate. That operator, evidently, does not correspond to any physical operation.

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 word *every* in (2), an important facet of the compatibility problem hinges on the *every* in proposition (1) above. 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 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 acceptable.

Several 'elementary' quantum theorems will now be stated and proved. The proofs are not new, but we shall nevertheless reproduce them in some detail in order to show their independence from strong correspondence. 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 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,

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

Proof (von Neumann, 1955, pp. 313–316; Gleason, 1957). Let $\{\varphi_n\}$ be a complete orthonormal set, and $A_{nk} \equiv \langle \varphi_n, A\varphi_k \rangle$. Since A is Hermitean, $A_{kn} = A^*_{nk}$, i.e., $\mathrm{Re}\,A_{kn} = \mathrm{Re}\,A_{nk}$, $\mathrm{Im}\,A_{kn} = -\mathrm{Im}\,A_{nk}$. Since $\{\varphi_n\}$ is complete,

$$\sum_n P_{\varphi_n} = 1$$

A may therefore be expressed as follows:

$$\begin{aligned} A &= \left(\sum_n P_{\varphi_n}\right) A \left(\sum_k P_{\varphi_k}\right) = \sum_{kn} P_{\varphi_n} A P_{\varphi_k} = \sum_{kn} |\varphi_n\rangle A_{nk} \langle\varphi_k| \\ &= \sum_n |\varphi_n\rangle A_{nn} \langle\varphi_n| \\ &\quad + \sum_{k<n} (|\varphi_n\rangle\langle\varphi_k| + |\varphi_k\rangle\langle\varphi_n|) \operatorname{Re} A_{nk} \\ &\quad + \sum_{k<n} i(|\varphi_n\rangle\langle\varphi_k| - |\varphi_k\rangle\langle\varphi_n|) \operatorname{Im} A_{nk} \end{aligned}$$

It is easily verified that the following operators are Hermitean:

$$\begin{aligned} P_{\varphi_n} &= |\varphi_n\rangle\langle\varphi_n| \\ K(n,k) &= |\varphi_n\rangle\langle\varphi_k| + |\varphi_k\rangle\langle\varphi_n| \\ \bar{K}(n,k) &= i(|\varphi_n\rangle\langle\varphi_k| - |\varphi_k\rangle\langle\varphi_n) \end{aligned}$$

Thus A now has the form of a linear combination of Hermitean operators:

$$A = \sum_n A_{nn} P_{\varphi_n} + \sum_{k<n} K(n,k) \operatorname{Re} A_{nk} + \sum_{k<n} \bar{K}(n,k) \operatorname{Im} A_{nk}$$

Now by P2 the mean value functional $m(A)$ is linear. Hence

$$\begin{aligned} m(A) &= \sum_n A_{nn} m(P_{\varphi_n}) + \sum_{k<n} [\operatorname{Re} A_{nk} m(K(n,k)) \\ &\quad + \operatorname{Im} A_{nk} m(\bar{K}(n,k))] \\ &= \sum_{kn} \rho_{nk} A_{nk} = \operatorname{Tr}(\rho A) \end{aligned}$$

where ρ is defined by its matrix elements:

$$\begin{aligned} \rho_{nn} &\equiv m(P_{\varphi_n}) \\ \rho_{kn} &\equiv \tfrac{1}{2} m(K(n,k)) - \tfrac{1}{2} i m(\bar{K}(n,k)), \qquad k < n \\ \rho_{nk} &\equiv \tfrac{1}{2} m(K(n,k)) + \tfrac{1}{2} i m(\bar{K}(n,k)), \qquad k < n \end{aligned}$$

Finally, by P2, $m(B)$ is real for any Hermitean B, in particular for P_{φ_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 *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 such 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 Hermitean operators to have discrete spectra; similar propositions hold for the continuous case.

Th2: 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.

Proof: $W_{\mathscr{A}}(a_k; \rho)$ is (by the physical definition of probability) the mean value of the observable $\mathscr{F}_k(\mathscr{A})$, where $\mathscr{F}_k$ is defined by

$$\mathscr{F}_k(a) = \begin{cases} 1, & a = a_k \\ 0, & a \neq a_k \end{cases}$$

By P1, $\mathscr{F}_k(A) \leftrightarrow \mathscr{F}_k(\mathscr{A})$. Hence, by Th1,

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

Consider the spectral expansion‡ of A and form $\mathscr{F}_k(A)$:

$$A = \sum_{nd_n} a_n P_{\alpha_{nd_n}}$$

$$\mathscr{F}_k(A) = \sum_{nd_n} \mathscr{F}_k(a_n) P_{\alpha_{nd_n}} = \sum_{d_k} P_{\alpha_{kd_k}} = P_{\mathscr{H}_k}$$

Therefore

$$W_{\mathscr{A}}(a_k; \rho) = \mathrm{Tr}\,(\rho P_{\mathscr{H}_k}) \qquad \text{Q.E.D.}$$

† In general, *open* 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 *indivisible* [compare Park, 1968a)].

‡ 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 α_{nd_n}, $d_n = 1, 2, \ldots$.

Th3: $\mathrm{Tr}\,\rho = 1$.

Proof: Consider the trivial observable $\mathscr{I}$ defined as follows: measure any observable $\mathscr{A}$; whatever the result, we shall say that an $\mathscr{I}$-measurement has been performed with numerical result unity. Symbolically, $\mathscr{I} = \mathscr{F}(\mathscr{A})$ where $\mathscr{F}(a) = 1$. The operator corresponding to $\mathscr{I}$ is therefore

$$\mathscr{F}(A) = \sum_{kd_k} \mathscr{F}(a_k) P_{\alpha_{kd_k}} = \sum_{kd_k} P_{\alpha_{kd_k}} = 1$$

the identity operator.

Now, it is obvious from the definition of $\mathscr{I}$ that $\langle \mathscr{I} \rangle = 1$; hence $\mathrm{Tr}(\rho 1) = \mathrm{Tr}\,\rho = 1$. Q.E.D.

Th4: The only possible results of $\mathscr{A}$-measurements are the eigenvalues $\{a_k\}$ of A, where $A \leftrightarrow \mathscr{A}$.

Proof: The probability Z that an $\mathscr{A}$-measurement will yield a number which is *not* an A-eigenvalue is equal to the mean value of the observable $\mathscr{G}(\mathscr{A})$, where

$$\mathscr{G}(a) = \begin{cases} 1, & a \text{ not an } A\text{-eigenvalue} \\ 0, & a \text{ an } A\text{-eigenvalue} \end{cases}$$

The operator to which $\mathscr{G}(\mathscr{A})$ corresponds is

$$\mathscr{G}(A) = \sum_n \mathscr{G}(a_n) P_{\mathscr{H}_n} = \sum_n (0) P_{\mathscr{H}_n} = 0$$

Hence $Z = \mathrm{Tr}(\rho 0) = 0$. 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, $P_\varphi \leftrightarrow \mathscr{P}$, has eigenvalues 0, 1 ($\varphi$ arbitrary). Thus by Th4‡ $\langle \mathscr{P} \rangle \geqslant 0$. By Th1, $\langle \mathscr{P} \rangle = \mathrm{Tr}(\rho P_\varphi)$. Hence for all φ, $\mathrm{Tr}(\rho P_\varphi) = \langle \varphi, \rho\varphi \rangle \geqslant 0$, 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 having complete orthonormal sets of eigenvectors. If $A \leftrightarrow \mathscr{A}$, then $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$.

† 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., $\mathscr{M}_1(\mathscr{A})$. 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 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 (Section 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 $\mathscr{F}(\mathscr{A},\mathscr{B})$ may be operationally defined as follows: measure $\mathscr{A}$ and $\mathscr{B}$ simultaneously, substitute the results a, 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 us directly to an old and interesting quantum problem (Groenewold, 1946; Shewell, 1959): given the correspondences $A\leftrightarrow\mathscr{A}$, $B\leftrightarrow\mathscr{B}$, ... and a compound observable $\mathscr{F}(\mathscr{A},\mathscr{B}\ldots)$, what F corresponds to $\mathscr{F}$? Note that if P1S is adopted, the *existence* of such an F is assured (if $\mathscr{A},\mathscr{B},\ldots$ 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 $F\leftrightarrow\mathscr{F}(\mathscr{A},\mathscr{B},\ldots)$ 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 $\mathscr{F}$, 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 $\mathscr{E}(A)$ and $\mathscr{N}(\mathscr{F})$ be defined as follows: $\mathscr{E}(A)$ is the set of eigenvalues belonging to the operator A; $\mathscr{N}(\mathscr{F})$ is the set of conceivable measurement results associated with an observable $\mathscr{F}$. When $\mathscr{F}=\mathscr{A}$, $\mathscr{N}(\mathscr{A})=\mathscr{E}(A)$ by Th4. However, when $\mathscr{F}$ is a function of $\mathscr{A}$ and $\mathscr{B}$, for example, 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 such a case, $\mathscr{E}(F) \subset \mathscr{N}(\mathscr{F})$. 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 now be expressed as follows: if $F \leftrightarrow \mathscr{F}(\mathscr{A}, \mathscr{B}, \ldots)$, then

$$(c_1) \sum_{kl} W(a_k, b_l, \ldots; \rho)\, \mathscr{F}(a_k, b_l, \ldots) = \mathrm{Tr}\,(\rho F),$$

for every ρ

$$(c_2) \quad \mathscr{E}(F) \subseteq \mathscr{N}[\mathscr{F}(\mathscr{A}, \mathscr{B}, \ldots)]$$

It is easy to see that (c_1) arises from Th1 and the definition of $\mathscr{F}$ 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 (compare Sections 8–10). Nevertheless, condition (c_1) is not so mute as it appears, since for the proper choice of $\mathscr{F}$, it becomes independent of the form of W (compare Section 4).

It will be noted that both P1 and P1S include explicit postulation of the correspondence $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$, 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 $W_{\mathscr{A}}$) 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 $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$ 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 $A \leftrightarrow \mathscr{A}$.

Proof: 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

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

which implies

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

(Note that (c_2) is also satisfied.) Q.E.D.

The $\mathscr{F}(\mathscr{A})$ which has probably received more attention than any other 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 just the arithmetic mean of $\mathscr{M}_1(\mathscr{A})$'s on the ensemble of interest. Using P2, we may write $\mathscr{F}(\mathscr{A}) = (\mathscr{A} - m(A))^2$; then by the correspondence rule in P1, $\mathscr{F}(\mathscr{A}) \leftrightarrow (A - m(A)\,1)^2$. By definition,

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

$\Delta\mathscr{A}$, called the standard deviation, is a common statistical quantity measured in the obvious way as a function of measurement results from an ensemble.

Historically, $\Delta\mathscr{A}$ has often been linked erroneously to the problem of compatibility by way of 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,

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

where A, B are Hermitean operators representing quantum observables $\mathscr{A}$, $\mathscr{B}$ and $\Delta\mathscr{A}$, $\Delta\mathscr{B}$ signify the standard deviations for collectives of $\mathscr{A}$- and $\mathscr{B}$-measurements.

Hence this remarkable theorem has the following physical significance: Given two identical ensembles of quantum systems, if $\mathscr{A}$-measurements are performed on one ensemble, $\mathscr{B}$-measurements on another, then the quantities $\Delta\mathscr{A}$, $\Delta\mathscr{B}$, calculated from the measurement results, will satisfy the relation above.

The principal point here stressed 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 $\Delta\mathscr{A}\,\Delta\mathscr{B}$, as presented above, is not even calculated from *simultaneous* measurements of $\mathscr{A}$ and $\mathscr{B}$ 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 noncommuting observables. For the interesting special case of position $\mathscr{X}$ and momentum $\mathscr{P}$ (whose operators satisfy $[X,P] = i\hbar 1$) this interpretation typically runs as follows: It is impossible to measure simultaneously $\mathscr{X}$ and $\mathscr{P}$ *exactly*; the product of the 'inaccuracies' $\Delta\mathscr{X}\,\Delta\mathscr{P}$ is never less than what the uncertainty principle allows, i.e., $\Delta\mathscr{X}\,\Delta\mathscr{P} \geqslant \frac{1}{2}\hbar$. 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 $\mathscr{X}$ and $\mathscr{P}$ exactly' here refers experimentally to a collective of $\mathscr{X}$- and $\mathscr{P}$-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 $\mathscr{X}$-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 $\Delta\mathscr{X} = 0$.

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 $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$ may be used to prove that any two operators jointly measurable in this trivial sense necessarily commute. To see this, assume the existence of an observable $\mathscr{C}$ and functions $\mathscr{F}$, $\mathscr{G}$ such that the observables $\mathscr{A}$, $\mathscr{B}$ are expressible in the form

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

By the above correspondence rule, if $C \leftrightarrow \mathscr{C}$, $C\gamma_k = c_k\gamma_k$,

$$\mathscr{F}(\mathscr{C}) \leftrightarrow \mathscr{F}(C) = \sum_k \mathscr{F}(c_k) P_{\gamma_k}$$

$$\mathscr{G}(\mathscr{C}) \leftrightarrow \mathscr{G}(C) = \sum_n \mathscr{G}(c_n) P_{\gamma_n}$$

Since two different operators cannot correspond to the same observable,† it follows that $A = \mathscr{F}(C)$, $B = \mathscr{G}(C)$. Hence

$$[A, B] = \sum_{kn} \mathscr{F}(c_k)\,\mathscr{G}(c_n)[P_{\gamma_k}, P_{\gamma_n}] = 0$$

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 $\mathscr{A}$ and $\mathscr{B}$, and (2) the only condition under which $\Delta\mathscr{A}\,\Delta\mathscr{B} = 0$ may hold, it is sometimes claimed via a misinterpretation of the uncertainty principle that the only simultaneous measurements per-

† If $\mathscr{A} \leftrightarrow A_1$ and also $\mathscr{A} \leftrightarrow A_2$, then for every ρ, $\langle \mathscr{A} \rangle = \mathrm{Tr}(\rho A_1) = \mathrm{Tr}(\rho A_2)$; but this implies $A_1 = A_2$.

mitted 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 $\mathscr{A}$ and $\mathscr{B}$ are simultaneously measurable through an auxiliary observable $\mathscr{C}$, then joint probabilities for the results of $\mathscr{M}_1(\mathscr{A})$ and $\mathscr{M}_1(\mathscr{B})$ are calculable through single probabilities associated with $\mathscr{M}_1(\mathscr{C})$. Given A, B satisfying the necessary condition $[A,B]=0$, the problem is to find $\mathscr{F}$, $\mathscr{G}$ and $\mathscr{C} \leftrightarrow C$ such that

$$\mathscr{A} \leftrightarrow A = \mathscr{F}(C) \leftrightarrow \mathscr{F}(\mathscr{C})$$

$$\mathscr{B} \leftrightarrow B = \mathscr{G}(C) \leftrightarrow \mathscr{G}(\mathscr{C})$$

By a mathematical theorem generally omitted from standard quantum textbooks, the condition $[A,B]=0$ is sufficient to insure the existence of $\mathscr{F}$, $\mathscr{G}$, and $\mathscr{C}$ such that

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

While some cases of the general theorem (von Neumann, 1931) 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 ε
by the relations

$$a_n = \mathscr{F}(c_{nl}), \qquad b_m = \mathscr{G}(c_{km})$$

This determines the desired C; its spectral expansic

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

Hence

$$\mathscr{F}(C) = \sum_{kl} \mathscr{F}(c_{kl}) P_{\gamma_{kl}} = \sum_{kl} a_k P_{\gamma_{kl}} = A$$

$$\mathscr{G}(C) = \sum_{kl} \mathscr{G}(c_{kl}) P_{\gamma_{kl}} = \sum_{kl} b_l P_{\gamma_{kl}} = B$$

(Note that $\mathscr{F}$, $\mathscr{G}$, and C are not unique.)

To find the joint probability $W(a_k, b_l; \rho)$ that $\mathscr{M}_1(\mathscr{A})$ and $\mathscr{M}_1(\mathscr{B})$ will yield (a_k, b_l) for the state ρ, we simply calculate $W_{\mathscr{C}}(c_{kl}; \rho)$, the probability that $\mathscr{M}_1(\mathscr{C})$ will yield c_{kl}, by the standard quantum mechanical rule for single probabilities:

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

The non-uniqueness of $\mathscr{F}$, $\mathscr{G}$, and C presents no problem since $W(a_k, b_l; \rho)$ 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 $\varphi(x, y)$. The operators corresponding to co-ordinates $\mathscr{X}$ and $\mathscr{Y}$ are postulated to be multiplication by x and y, respectively; but it is not necessary to postulate further that $\varphi^* \varphi$ is the joint probability density for $\mathscr{M}_1(\mathscr{X})$ and $\mathscr{M}_1(\mathscr{Y})$. Since $[x, y] = 0$, we may conclude that $\mathscr{X}$ and $\mathscr{Y}$ are jointly measurable in the trivial sense through an auxiliary observable $\mathscr{C} \leftrightarrow C$; 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 $\mathscr{X}$ and $\mathscr{Y}$ in the (pure) state $\rho = P_\varphi$ is given by

$$\begin{aligned} w(x_1, y_1; P_\varphi) &= w_{\mathscr{C}}(c_{x_1 y_1}; P_\varphi) = |\langle \gamma_{x_1 y_1}, \varphi \rangle|^2 \\ &= \left| \int dx \int dy\, \delta(x - x_1)\, \delta(y - y_1)\, \varphi(x, y) \right|^2 \\ &= \varphi^*(x_1, y_1)\, \varphi(x_1, y_1) \end{aligned}$$

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 $\mathscr{X}$ and $\mathscr{Y}$ for the impinging particle may

be made by reading the $\mathscr{X}$ and $\mathscr{Y}$ scales separately. Two numbers thus emerge, the $\mathscr{X}$- and $\mathscr{Y}$-measurements having been performed non-trivially. To construct an apparatus which simultaneously measures $\mathscr{X}$ and $\mathscr{Y}$ in the trivial sense, we require an observable $\mathscr{C}$ and the functions $\mathscr{F}$, $\mathscr{G}$. A suitable observable $\mathscr{C}$ 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 $\mathscr{C}$-measurement upon the particle.

If the functions $\mathscr{F}$ and $\mathscr{G}$ 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 $\mathscr{X}$ and $\mathscr{Y}$.

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 $\mathscr{A}$ and $\mathscr{B}$ simultaneously in the trivial sense, i.e., if $[A,B]=0$, then the joint probability $W(a_k,b_l;\rho)$ calculated using an auxiliary observable $\mathscr{C}$ is valid for all simultaneous $\mathscr{A}$, $\mathscr{B}$-measurements regardless of whether or not they are actually performed by means of $\mathscr{C}$-measurements. 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 $W(a_k,b_l;\rho)$, $[A,B]=0$, 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 $\mathscr{A}$ 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_2) still hold? Is the

† 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, namely, that of the continuum.

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 $\mathscr{A}$ must each yield the same result. Formally, we may choose $\mathscr{C} = \mathscr{A}$, $\mathscr{F}(\mathscr{C}) = \mathscr{G}(\mathscr{C}) = \mathscr{A}$. Assignment of probabilities to $\mathscr{F}(\mathscr{C})$- and $\mathscr{G}(\mathscr{C})$-measurements via ordinary quantal analysis of $\mathscr{C}$-measurements obviously leads to nonzero probability only when 'both' $\mathscr{A}$-measurements yield identical results, namely, that of the auxiliary $\mathscr{C}(=\mathscr{A})$-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 (von Neumann, 1955, pp. 255–230) 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 (Mackey, 1963; Segal, 1947; Varadarajan, 1962) 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 $\mathscr{A}$ and $\mathscr{B}$. Assume that $\mathscr{A}$ and $\mathscr{B}$ are compatible (jointly measurable). If $\mathscr{A}$ and $\mathscr{B}$ 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 $\mathscr{S}(\mathscr{D})$ expressed symbolically by

$$\mathscr{S} = \mathscr{A} + \mathscr{B} \qquad (\mathscr{D} = \mathscr{A} - \mathscr{B})$$

Clearly if $\mathscr{A}$ and $\mathscr{B}$ are compatible observables, $\mathscr{S}$ and $\mathscr{D}$ are observables. Therefore, in accordance with a widely accepted quantal

postulate (P1S) there must exist Hermitean operators S and D representing observables $\mathscr{S}$ and $\mathscr{D}$. 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:

$$(c_1) \sum_{kl} W(a_k, b_l; \rho)\, \mathscr{F}(a_k, b_l) = \mathrm{Tr}\,(\rho S), \qquad \text{for every } \rho$$

Here

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

therefore

$$\sum_k \left[\sum_l W(a_k, b_l; \rho)\right] a_k + \sum_l \left[\sum_k W(a_k, b_l; \rho)\right] b_l = \mathrm{Tr}\,(\rho S)$$

Because of the additive form of $\mathscr{F}$, 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:

$$\sum_k W(a_k, b_l; \rho) = W_{\mathscr{B}}(b_l; \rho) = \mathrm{Tr}\,(\rho P_{\beta_l})$$

$$\sum_l W(a_k, b_l; \rho) = W_{\mathscr{A}}(a_k; \rho) = \mathrm{Tr}\,(\rho P_{\alpha_k})$$

Hence, for every ρ,

$$\sum_k \mathrm{Tr}\,(\rho P_{\alpha_k})\, a_k + \sum_l \mathrm{Tr}\,(\rho P_{\beta_l})\, b_l = \mathrm{Tr}\,(\rho S)$$

This determines S uniquely, as follows:

$$\begin{aligned} \mathrm{Tr}\,(\rho S) &= \mathrm{Tr}\left(\rho \sum_k a_k P_{\alpha_k}\right) + \mathrm{Tr}\left(\rho \sum_l b_l P_{\beta_l}\right) \\ &= \mathrm{Tr}\,(\rho A) + \mathrm{Tr}\,(\rho B) \\ &= \mathrm{Tr}\,[\rho(A + B)], \qquad \text{for every } \rho. \end{aligned}$$

Therefore

$$S = A + B$$

Similarly,

$$D = A - B$$

From the correspondence rule $\mathscr{F}(\mathscr{A}) \leftrightarrow \mathscr{F}(A)$ of Section 2 together with those just derived there follows another:

$$\mathscr{A}\mathscr{B} \leftrightarrow \tfrac{1}{2}(AB + BA)$$

where $\mathscr{A}\mathscr{B}$ denotes the observable measured by multiplying the results of simultaneous measurements of $\mathscr{A}$ and $\mathscr{B}$. The argument proceeds as follows. If $\mathscr{A}$ and $\mathscr{B}$ are compatible, $\mathscr{S}$ and $\mathscr{D}$ are also, as well as

$\mathscr{S}^2$ and $\mathscr{D}^2$. Thus $\mathscr{R} \equiv \frac{1}{4}(\mathscr{S}^2 - \mathscr{D}^2)$ 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 $\mathscr{A}$ and $\mathscr{B}$ 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 $\mathscr{R}$-measurement: $r = \frac{1}{4}[(a+b)^2 - (a-b)^2] = ab$. Hence, the operator for the product $\mathscr{A}\mathscr{B}$ 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 $\mathscr{A}$ and $\mathscr{B}$ are compatible and $\mathscr{A} \leftrightarrow A$, $\mathscr{B} \leftrightarrow B$, then

$$[A, B] = 0$$

Proof: Since $\mathscr{A}$ and $\mathscr{B}$ are compatible, it follows that any function of $\mathscr{A}$ and $\mathscr{B}$ is observable (simply by measuring $\mathscr{A}$ and $\mathscr{B}$ simultaneously and using the results to evaluate the function); in particular, consider the function $\mathscr{A}^2\mathscr{B}$. By the correspondence rules discussed above, $\mathscr{A}^2 \leftrightarrow A^2$ and hence

$$\mathscr{A}^2\mathscr{B} \leftrightarrow \frac{A^2B + BA^2}{2} \equiv P_1$$

But the product correspondence rule also implies

$$\mathscr{A}^2\mathscr{B} = \mathscr{A}(\mathscr{A}\mathscr{B}) \leftrightarrow \frac{1}{2}\left(A\frac{AB+BA}{2} + \frac{AB+BA}{2}A\right)$$

$$= \tfrac{1}{4}(A^2B + 2ABA + BA^2) \equiv P_2$$

Since quantum theory cannot tolerate the ambiguity of having several operators for a single observable (compare footnote, p. 230), we must set $P_1 = P_2$ and accept the consequences. The resultant equation is

$$ABA = \tfrac{1}{2}(A^2B + BA^2)$$

Moreover, the same argument is applicable to any functions $\mathscr{F}(\mathscr{A})$ and $\mathscr{G}(\mathscr{B})$; thus

$$\mathscr{F}(A)\,\mathscr{G}(B)\,\mathscr{F}(A) = \tfrac{1}{2}[\mathscr{F}^2(A)\,\mathscr{G}(B) + \mathscr{G}(B)\,\mathscr{F}^2(A)] \quad (4.1)$$

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 $\mathscr{F}_k$ and $\mathscr{G}_l$ as follows:

$$\mathscr{F}_k(a) = \delta_{aa_k}, \qquad \mathscr{G}_l(b) = \delta_{bb_l}$$

Then

$$\mathscr{F}_k(\mathscr{A}) \leftrightarrow \mathscr{F}_k(A) = \sum_n \delta_{a_n} a_k P_{\alpha_n} = P_{\alpha_k}$$

$$\mathscr{G}_l(\mathscr{B}) \leftrightarrow \mathscr{G}_l(B) = \sum_m \delta_{b_m} b_l P_{\beta_m} = P_{\beta_l}$$

Substituting these operators into condition (4.1), we have for each k, l,

$$P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \tfrac{1}{2}(P_{\alpha_k}^2 P_{\beta_l} + P_{\beta_l} P_{\alpha_k}^2)$$

But

$$P_{\alpha_k}^2 = P_{\alpha_k}, \qquad P_{\beta_l}^2 = P_{\beta_l}$$

Hence

$$P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \tfrac{1}{2}(P_{\alpha_k} P_{\beta_l} + P_{\beta_l} P_{\alpha_k}) \tag{4.2}$$

Multiply (4.2) on the left by P_{α_k}:

$$P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \tfrac{1}{2}(P_{\alpha_k} P_{\beta_l} + P_{\alpha_k} P_{\beta_l} P_{\alpha_k}) \tag{4.3}$$

Multiply (4.2) on the right by P_{α_k}:

$$P_{\alpha_k} P_{\beta_l} P_{\alpha_k} = \tfrac{1}{2}(P_{\alpha_k} P_{\beta_l} P_{\alpha_k} + P_{\beta_l} P_{\alpha_k}) \tag{4.4}$$

Subtract (4.4) from (4.3):

$$P_{\alpha_k} P_{\beta_l} - P_{\beta_l} P_{\alpha_k} = 0$$

Therefore

$$[A, B] = \sum_{kl} a_k b_l [P_{\alpha_k}, P_{\beta_l}] = 0$$

This completes the proof of von Neumann's simultaneous measurability theorem.

Expressed succinctly, it says that if $\mathscr{A}$ and $\mathscr{B}$ 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 (Birkhoff & von Neumann, 1936; Piron, 1964; Reichenbach, 1944)

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

$$\mathscr{A} \leftrightarrow A = \sum_k a_k P_{\alpha_k}$$

and the proposition $\mathscr{P}_n$: '$\mathscr{M}_1(\mathscr{A})$ would yield a_n.' The proposition $\mathscr{P}_n$ is just the observable measured as follows: measure $\mathscr{A}$; if a_n results, assign to $\mathscr{P}_n$ the value 1 (by convention); if a_k $(\neq a_n)$ results, 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 $\mathscr{F}_n(a_k) = \delta_{a_n a_k}$. Hence

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

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 non-commuting observables must of course 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 relations $\mathscr{P}$ 'or' $\mathscr{Q}$ $\mathscr{R} \equiv \mathscr{P} \cup \mathscr{Q}$ and $\mathscr{P}$ 'and' $\mathscr{Q}$ $\mathscr{R} \equiv \mathscr{P} \cap \mathscr{Q}$:

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

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

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 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. To see this, note that

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

but $(\mathscr{P} \cap \mathscr{Q}) \cup (\mathscr{P} \cap \mathscr{R})$ is an absurd proposition because 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 '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 construct such counterexamples.

It is instructive to consider first a typical quantum theory of measurement for a single observable, namely, the 'time-of-flight' method for measuring the momentum $\mathscr{P}$ of an electron. We assume that the rule of correspondence for position $\mathscr{X}$ 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 $w_{\mathscr{P}}(p;\psi)$ for $\mathscr{M}_1(\mathscr{P})$ at the time of preparation is easily calculated:

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

where

$$\mathscr{P} \leftrightarrow P = \frac{\hbar}{i} \frac{\partial}{\partial x}, \qquad \pi_p = (2\pi\hbar)^{-1/2} \exp \frac{ipx}{\hbar},$$

$$P\pi_p = p\pi_p, \qquad \langle \pi_{p_1}, \pi_{p_2} \rangle = \delta(p_1 - p_2)$$

This distribution is the quantum mechanical touchstone for deciding whether or not a proposed experiment which generates numbers via the established operational definition for $\mathscr{M}_1(\mathscr{X})$ 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 (Feynmann, 1965) is simple: wait a very long time $(t \to \infty)$ as the electron moves *freely*, then measure the observable $\mathscr{F}(\mathscr{X}) = m\mathscr{X}/t$, where m is the electron mass. The resultant number then counts as the result of $\mathscr{M}_1(\mathscr{P})$ at $t = 0$. To justify this operational definition of $\mathscr{P}$ quantum mechanically, we 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 $mx/t \in (p_1, p_2)$ as $t \to \infty$. In short, it must be shown that

$$W_{\mathscr{P}}[p \in (p_1, p_2); \psi(x, 0)] = W_{\mathscr{F}(\mathscr{X})}\left[\frac{mx}{t} \in (p_1, p_2); \psi(x, t), t \to \infty\right]$$

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

$$\psi(x,t) = \sqrt{\left(\frac{m}{2\pi i\hbar t}\right)} \exp\frac{imx^2}{2\hbar t} \int_{-\infty}^{\infty} \exp\left[-\frac{im}{\hbar t}\left(xx_1 - \frac{x_1^2}{2}\right)\right]\psi(x_1, 0)\, dx_1$$

Now,

$$\begin{aligned} W_{\mathscr{F}(\mathscr{X})}&\left[\frac{mx}{t} \in (p_1, p_2); \psi(x,t)\right] \\ &= W_{\mathscr{X}}\left[x \in \left(\frac{p_1 t}{m}, \frac{p_2 t}{m}\right); \psi(x,t)\right] \\ &= \int_{p_1 t/m}^{p_2 t/m} |\psi(x,t)|^2\, dx \\ &= \frac{m}{2\pi\hbar t} \int_{p_1 t/m}^{p_2 t/m} dx \left| \int_{-\infty}^{\infty} \exp\left[-\frac{im}{\hbar t}\left(xx_1 - \frac{x_1^2}{2}\right)\right]\psi(x_1)\, dx_1 \right|^2 \end{aligned}$$

Changing the variable of integration by letting $x = pt/m$, we obtain

$$W_{\mathscr{F}(\mathscr{X})}\left[\frac{mx}{t} \in (p_1, p_2); \psi(x,t)\right] = \int_{p_1}^{p_2} \frac{dp}{2\pi\hbar} \left| \int_{-\infty}^{\infty} \exp\left[\frac{im}{2\hbar t} x_1^2 - \frac{ipx_1}{\hbar}\right] \psi(x_1)\, dx_1 \right|^2$$

Since $\psi(x_1) = 0$ for $x_1 \notin (-x_0, x_0)$, the integral over x_1 is 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

$$W_{\mathscr{F}(\mathscr{X})}\left[\frac{mx}{t} \in (p_1, p_2); \psi(x,t), t \to \infty\right] = \int_{p_1}^{p_2} \left| \int_{-\infty}^{\infty} \frac{\exp(-ipx_1/\hbar)}{\sqrt{(2\pi\hbar)}} \psi(x_1)\, dx_1 \right|^2 dp$$

But this equals

$$\int_{p_1}^{p_2} w_{\mathscr{P}}[p; \psi(x,0)]\, dp = W_{\mathscr{P}}[p \in (p_1, p_2); \psi(x,0)]$$

Hence 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 thus fully certified quantum mechanically as an operational definition of $\mathscr{P}$ (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 $\mathscr{P}$-measurements is required. Indeed further *quantal* 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 $t \to \infty$, $p \to mx/t$. 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 $\mathscr{M}_1(\mathscr{P})$ at $t = 0$ would yield also determines the result that $\mathscr{M}_1(\mathscr{P})$ 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,

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

Thus by the same reasoning which validated the time-of-flight method as a rule of correspondence for $\mathscr{M}_1(\mathscr{P})$ at $t = 0$, we can likewise 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 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 $\mathscr{M}_1(\mathscr{P})$ would have yielded $\mathscr{F}(x)$ where x is the result of the $\mathscr{M}_1(\mathscr{X})$. Hence, contrary to the prohibitions of von Neumann's theorem, we have an empirical method for the simultaneous measurement of $\mathscr{X}$ *and* $\mathscr{P}$, 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 $\mathscr{P}$-measurement. Although it is true even classically that time-of-flight $\mathscr{P}$-measurements are never perfect for finite t (unless $\mathscr{X}$ 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 $(\mathscr{X},\mathscr{P})$-measurements in general; for example, a magnetic deflection method for joint measurement of $\mathscr{X}$ and $\mathscr{P}$ which does not require $t \to \infty$ will be outlined later (Section 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 $\mathscr{X}$ and $\mathscr{P}$ for example, cannot be measured simultaneously. The usual argument seems to have been originated by Heisenberg and may be summarized by his statement (Heisenberg, 1930, p. 20) that 'the uncertainty relation does not refer to the past'. In the time-of-flight experiment, for example, 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 (1930) such 'know-

ledge 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 response to this distinctly philosophical argument, we offer the following comments:

(1) *Knowledge* is an ambiguous word in discussions regarding quantum measurement. From a strict quantal point of view an electron *never possesses properties* $\mathscr{X}$, $\mathscr{P}$, 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 $\mathscr{X},\mathscr{P}$-values from $\mathscr{X},\mathscr{P}$-measurements; the relation $\Delta\mathscr{X}\,\Delta\mathscr{P} \geqslant \hbar/2$ 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 be to 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 $\mathscr{X},\mathscr{P}$-values upon measurement in no way violates the uncertainty principle.

(3) In the time-of-flight method, the $\mathscr{X},\mathscr{P}$-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 $\mathscr{X}$-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 $\mathscr{X},\mathscr{P}$-values refer to a past time is no special property of *simultaneous* measurements; it is characteristic of all quantum measurements. 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 $\psi(x, 0)$. Indeed, if the physical significance of $\mathscr{X}, \mathscr{P}$-values is a matter of 'personal belief', then all measurement results for *single* observables are likewise of solipsistic value only.

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

The time-of-flight method for $\mathscr{M}_2(\mathscr{P})$ exhibits a curious feature which seems at first paradoxical. Only $\mathscr{X}$ is 'directly' measured; $\mathscr{P}$ is then measured by calculation of $\mathscr{F}(\mathscr{X})$. Yet if $\mathscr{P} = \mathscr{F}(\mathscr{X})$, it follows (compare Section 2) for the respective operators that $P = \mathscr{F}(X)$, 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 $\mathscr{X}_0 \leftrightarrow X$, $\mathscr{P}_0 \leftrightarrow P$, $\mathscr{X}_t \leftrightarrow X_t$, $\mathscr{P}_t \leftrightarrow P_t$, $X_t = T^\dagger X T$, $P_t = T^\dagger P T$, T being the *free* evolution operator. In these terms, it would appear that the time-of-flight scheme is based on the relation $P = (m/t) X_t$, $t \to \infty$. But from this expression and momentum conservation ($P = P_t$), it follows that

$$[X, P] = [T^\dagger X T, T^\dagger P T] = [T^\dagger X T, P] = \left[X_t, \frac{m X_t}{t}\right] = 0,$$

which contradicts $[X, P] = i\hbar 1$,

The resolution to this dilemma provides interesting information regarding quantum measurement. The error in the above reasoning inheres in the assumption that $P = m X_t/t$ just because $\mathscr{P} = m\mathscr{X}_t/t$ is used in the time-of-flight arrangement. Actually, the relation $\mathscr{P} = m\mathscr{X}_t/t$ holds only for initial wave functions $\psi(x, 0)$ of compact support that develop freely until the measurement of $\mathscr{X}_t$, $t \to \infty$. In general, $[X, P] \neq 0$ and no $\mathscr{F}$ exists such that $\mathscr{P} = \mathscr{F}(\mathscr{X})$. Nevertheless, this analysis reveals two interesting points concerning quantum measurements: (1) Even if $[\mathscr{A}, \mathscr{B}] \neq 0$ so that no *general* trivial joint measurement scheme for $\mathscr{A}$ and $\mathscr{B}$ can be constructed, it may still be possible, for certain states and measurement arrangements, to measure $\mathscr{B}$ as a function of $\mathscr{A}$ (or vice versa). (2) Conversely, the fact that $\mathscr{B} = \mathscr{F}(\mathscr{A})$ for some *particular* $\mathscr{M}_2(\mathscr{B})$ does not imply $B = \mathscr{F}(A)$. However, as shown in Section 2, if $\mathscr{B} = \mathscr{F}(\mathscr{A})$ always, then the operator relation $B = \mathscr{F}(A)$ 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 $\mathscr{A}_1$, $\mathscr{B}_1$ and $\mathscr{A}_2$ associated with S_1 and S_2, respectively. Suppose $[A_1, B_1] \neq 0$ and denote eigenvectors and eigenvalues as follows:

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

Let $S_1 + S_2$ be noninteracting but in a correlated state:

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

If $\mathscr{A}_2$ has an established operational definition, the correlation in Ψ which relates $\mathscr{M}_1(\mathscr{A}_1)$-results to $\mathscr{M}_1(\mathscr{A}_2)$-results may be exploited in the standard way to construct an $\mathscr{M}_2(\mathscr{A}_1)$. As in the time-of-flight $\mathscr{M}_2(\mathscr{P})$ case, 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 (compare Section 3). The joint probability $W(a_k^{(1)}, a_l^{(2)}; \Psi)$ is therefore easily calculated:

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

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 $\mathscr{A}_2$, 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 Axiomatic Root of Quantal Inconsistencies*

The necessary conclusions to be drawn from the last two sections may be summarized as follows: (1) The quantal postulates (P1S, 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 and 4) in order to derive a theorem about simultaneous measurement of several observables from axioms referring only to single observable measurements $\mathscr{M}_1(\mathscr{A})$, 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 $\mathscr{A} + \mathscr{B}$, the latter condition assumed a simple form and, moreover, it even sufficed to *determine* uniquely the correspondence $\mathscr{A} + \mathscr{B} \leftrightarrow A + B$ upon which the simultaneous measurability theorem was ultimately based. In fact, once this correspondence is established, the logic of the theorem cannot be impugned, as careful restudy of its proof (Section 4) will demonstrate.

We therefore direct attention to the correspondence $\mathscr{A} + \mathscr{B} \leftrightarrow A + B$. As just noted, condition (c_1) alone *implied* this rule. To be more explicit, P1S guaranteed the existence of an operator corresponding to the observable $\mathscr{A} + \mathscr{B}$; that operator would necessarily satisfy (c_1) and (c_2). It then turned out that (c_1) for $\mathscr{A} + \mathscr{B}$ could be satisfied by only *one* operator, $A + B$. *Thus condition* (c_2) *was never used.* This observation provides an important clue in our search for the false hypothesis which made possible the proof of von Neumann's (evidently false) simultaneous measurability theorem.

Is $\mathscr{A} + \mathscr{B}$ really an observable? If not, P1S cannot be invoked to assure the existence of an operator counterpart. To establish the observability of $\mathscr{A} + \mathscr{B}$, 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 $\mathscr{M}_1(\mathscr{A}_1)$ and $\mathscr{M}_1(\mathscr{B}_1)$. Since the experimenter is obviously free to add the two results, it is apparent that $\mathscr{A}_1 + \mathscr{B}_1$ is indeed observable. *Therefore, by P1S, there must exist an operator* S *such that* $\mathscr{A}_1 + \mathscr{B}_1 \leftrightarrow S$.

For simplicity, let system S_1 be a 'spin' whose relevant states and operators refer to a two-dimensional spinor space. For noncommuting observables $\mathscr{A}_1$, $\mathscr{B}_1$, take x- and z-components of spin $\mathscr{S}_x$ and $\mathscr{S}_z$. In terms of the Pauli operators, we have

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

Now, simultaneous measurements of $\mathscr{A}_1$ and $\mathscr{B}_1$ using the correlation scheme involving auxiliary system S_2 will by Th4 always yield one of these eigenvalue pairs: $(\hbar/2,\hbar/2)$, $\hbar/2,-\hbar/2)$, $(-\hbar/2,\hbar/2)$, $(-\hbar/2,-\hbar/2)$. (The two eigenvalues of any Pauli matrix are 1, -1.) Hence upon addition to obtain $\mathscr{A}_1+\mathscr{B}_1$-measurement results, only the three values, $\hbar$, 0, $-\hbar$ are possible. To use the set notation of Section 2,

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

and by consistency condition (c_2), if $\mathscr{A}_1+\mathscr{B}_1\leftrightarrow S$, quantum mechanics would be self-contradictory unless

$$(c_2)\qquad \mathscr{E}(S)\subseteq\mathscr{N}(\mathscr{A}_1+\mathscr{B}_1)$$

But (c_1) must also be satisfied by S and, as shown in Section 4, the *only* S meeting this requirement is, for the $\mathscr{A}_1+\mathscr{B}_1$ 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 $\hbar/\sqrt{2}$, $-\hbar/\sqrt{2}$; i.e., $\mathscr{E}(S)=\{-\hbar/\sqrt{2},\hbar\sqrt{2}\}$. Comparing the sets $\mathscr{N}(\mathscr{A}_1+\mathscr{B}_1)$ and $\mathscr{E}(S)$, $\mathscr{A}_1+\mathscr{B}_1\leftrightarrow S$, we find that

$$\mathscr{N}(\mathscr{A}_1+\mathscr{B}_1)\cap\mathscr{E}(S)=\phi$$

Thus the only operator S capable of satisfying (c_1) does not satisfy (c_2).

To summarize: $\mathscr{A}_1+\mathscr{B}_1$ is demonstrably observable. P1S then insures the existence of $S\leftrightarrow\mathscr{A}_1+\mathscr{B}_1$. 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 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, namely, $\mathrm{Tr}(\rho A)$. 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 justified.

If P1S is replaced by P1, what does von Neumann's theorem really prove? It becomes a *reductio ad absurdum* proof that the correspondence $\mathscr{A} + \mathscr{B} \leftrightarrow A + B$ and its consequence $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$ 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 $\mathscr{A}$ and $\mathscr{B}$, 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 $\mathscr{A} + \mathscr{B}$ or $\mathscr{A}\mathscr{B}$

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 (Temple, 1935; Peierls, 1935).

The basic premise is strong correspondence. As we have seen, it follows that if $A \leftrightarrow \mathscr{A}$, $B \leftrightarrow \mathscr{B}$ then

$$A + B \leftrightarrow \mathscr{A} + \mathscr{B}, \qquad \tfrac{1}{2}(AB + BA) \leftrightarrow \mathscr{A}\mathscr{B}$$

Now, when these rules together with $\mathscr{F}(A) \leftrightarrow \mathscr{F}(\mathscr{A})$ are applied to the observable $\mathscr{A}\mathscr{B}\mathscr{C}$, there results a threefold ambiguity:

$$(\mathscr{A}\mathscr{B})\mathscr{C} \leftrightarrow \tfrac{1}{2}[\tfrac{1}{2}(AB + BA)C + \tfrac{1}{2}C(AB + BA)]$$
$$= \tfrac{1}{4}(AB + BA)C + \tfrac{1}{4}C(AB + BA)$$

$$(\mathscr{C}\mathscr{A})\mathscr{B} \leftrightarrow \tfrac{1}{4}(CA + AC)B + \tfrac{1}{4}B(CA + AC)$$

$$(\mathscr{B}\mathscr{C})\mathscr{A} \leftrightarrow \tfrac{1}{4}(BC + CB)A + \tfrac{1}{4}A(BC + CB)$$

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 $\mathscr{A}\mathscr{B}\mathscr{C}$ 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 $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$ denoted *any* 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

† $[C,[A,B]] = 0$ 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

$$d(B,C)\,A + d(A,C)\,B = d\left(\frac{AB + BA}{2}, C\right)1$$

for every A, B, C. Hence

$$d(B,C) = d(A,C) = d\left(\frac{AB + BA}{2}, C\right) = 0$$

for every A, B, C.

regarded as erroneous, 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 erroneous; and the common root of both of these perplexing theorems, namely, 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 $\mathrm{Tr}(\rho A)$. 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 *are* 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 $\mathscr{M}_1(\mathscr{A})$ which give meaning to the primitive term *observable* as it appears in P1.) Thus none of the developments in linear algebra which form the theorctical structure of quantum theory will be affected by the elimination of strong correspondence. In fact a careful search

through quantum theory by the present writers 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. (For example, compare Theorems 1–5 in Section 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 $\mathscr{F}$ 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 $\mathscr{F}$, 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 $\mathscr{F}$, 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. Here is a summary of the correct interpretation of the theorems of this kind which were presented 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 'simultaneous measurability' theorem is merely a demonstration 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) $\mathscr{A}\mathscr{B}\mathscr{C} \leftrightarrow ?$: 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 $\mathscr{F}(\mathscr{C}) \leftrightarrow \mathscr{F}(C)$. To see this, recall that $[A,B] = 0$ implies the existence of C, $\mathscr{F}$, $\mathscr{G}$ such that $A = \mathscr{F}(C)$, $B = \mathscr{G}(C)$. Hence $\mathscr{A} + \mathscr{B} = \mathscr{F}(\mathscr{C}) + \mathscr{G}(\mathscr{C}) \leftrightarrow \mathscr{F}(C) + \mathscr{G}(C) = A + B$ and $\mathscr{A}\mathscr{B} = \mathscr{F}(\mathscr{C})\mathscr{G}(\mathscr{C}) \leftrightarrow \mathscr{F}(C)\mathscr{G}(C) = AB = BA$. In this connection it is instructive to consider a simple example which illustrates why consistency condition (c_2) required only $\mathscr{E}(F) \subseteq \mathscr{N}[\mathscr{F}(\mathscr{A},\mathscr{B})]$ rather than $\mathscr{E}(F) = \mathscr{N}[\mathscr{F}(\mathscr{A},\mathscr{B})]$. Let $\mathscr{A} = \mathscr{L}_z{}^2$, $\mathscr{B} = \hbar\mathscr{L}_z$, where $\mathscr{L}_z$ is the z-component of orbital momentum, $\mathscr{L}_z \leftrightarrow L_z = (\hbar/i)(\partial/\partial\varphi)$. Suppose $\mathscr{A}$ and $\mathscr{B}$ are measured simultaneously and the results added together. $\mathscr{N}(\mathscr{A} + \mathscr{B})$, the set of all *a priori* conceivable results of this procedure

is given by $\mathscr{N}(\mathscr{L}_z{}^2 + \hbar\mathscr{L}_z) = \{m^2\hbar^2 + n\hbar^2\}$, since $\mathscr{E}(L_z) = \{m\hbar\}$. Now, 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 $\mathscr{N}(\mathscr{A} + \mathscr{B})$, i.e.,

$$\mathscr{E}(L_z{}^2 + \hbar L_z) \subset \mathscr{N}(\mathscr{L}_z{}^2 + \hbar\mathscr{L}_z)$$

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 $\mathscr{L}_z{}^2$ 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 trivial joint measurement would involve simply measuring $\mathscr{L}_z$ and evaluating $\mathscr{L}_z{}^2 + \hbar\mathscr{L}_z$. Obviously, this procedure could yield only numbers in the set $\{k(k+1)\hbar^2\} = \mathscr{E}(L_z{}^2 + \hbar L_k)$. This demonstration merely affirms the consistency of (J_1) or (J_2) with the postulated correspondence $\mathscr{F}(\mathscr{C}) \leftrightarrow \mathscr{F}(C)$.

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 $\mathscr{A}$, $\mathscr{B}$, 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 *classical* harmonic oscillator, i.e., a system whose energy is given by the function

$$\mathscr{W}(x,p) = \frac{p^2}{2m} + \tfrac{1}{2}kx^2, \qquad k > 0$$

Supposedly, the oscillator is 'quantized' by using the established correspondences $\mathscr{X} \leftrightarrow X, \mathscr{P} \leftrightarrow P$ plus rules (1) and (2) to find operator H such that $\mathscr{W} \leftrightarrow H$. This procedure yields immediately

$$H = \frac{1}{2m}P^2 + \frac{k}{2}X^2$$

the energy operator which then serves to define the *quantum* harmonic oscillator. To see that this scheme is merely mnemotechnical and has no significant logical value, it is sufficient to note that the classical *function* $\mathscr{W}$ contributed nothing to the *operator* H except its functional form. If the definition of the quantum energy operator were based in a *physical* sense upon the classical energy function, then the operational definition of energy which $\mathscr{W}(x,p)$ entails would also be valid for the quantum energy operator: quantum energy would like classical energy be measurable by first measuring observables $\mathscr{X}, \mathscr{P}$, then using the results, x, p to evaluate $\mathscr{W}(x,p)$. Thus the quantum

energy operator H would be defined such that $\mathscr{W}(\mathscr{X},\mathscr{P}) \leftrightarrow H$, 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 $\mathscr{H}$ corresponds to the operator above, the form of which can be *remembered* by classical analogy. $\mathscr{H}$ has its own rules of correspondence, but among these the classical prescription—measure $\mathscr{X}$ and $\mathscr{P}$, evaluate $\mathscr{W}(x,p)$— *does not appear.* It is quite possible to measure $\mathscr{X}$ and $\mathscr{P}$ simultaneously and obtain thereby a value for $\mathscr{W}(x,p)$ 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 $\mathscr{M}_1(\mathscr{P})$ on S_1. Consider the S_1-subensemble defined by the property that such $\mathscr{M}_1(\mathscr{P})$'s would have yielded $p = p_1$; this subensemble is therefore characterized by the eigenfunction $\pi_p = (2\pi\hbar)^{-1}\exp(ip_1x/\hbar)$ (δ-function normalized), which obviously assigns equal relative probability to *all* results of $\mathscr{M}_1(\mathscr{X})$. Hence this kind of simultaneous $\mathscr{X},\mathscr{P}$-measurements upon the oscillator will often lead to values of $\mathscr{W}(x,p)$ which are not in the set $\mathscr{E}(H) = \{(n+\frac{1}{2})\hbar\sqrt{(k/m)}\}$. Hence by Th4 the classical energy function $\mathscr{W}(x,p)$ has in a sense nothing to do with the quantum energy observable $\mathscr{H} \leftrightarrow H$.

Let $w(x,p)$ be the joint probability density associated with the simultaneous measurements of $\mathscr{X}$ and $\mathscr{P}$. Our analysis of the relation between $\mathscr{W}(x,p)$ and H indicates clearly that we should *not* expect $w(x,p)$ to satisfy the following equation:†

$$\int dx \int dp w(x,p)\,\mathscr{W}(x,p) = \mathrm{Tr}\,(\rho H)$$

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

† Another example to illustrate this point may be constructed by contrasting measurements of the observable $\mathscr{H} + \frac{1}{2}\mathscr{X}^2$ 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 $\{(n+\frac{1}{2})\hbar\sqrt{[(k+1)/m]}\}$. Thus the operator $H + \frac{1}{2}X^2$ and the observable $\mathscr{H} + \frac{1}{2}\mathscr{X}^2$ do not refer to the same physical situation.

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 $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$ is false unless $[A,B] = 0$, or in other words, a proof that $[A,B] = 0$ is a necessary condition for the validity of $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$.

Hence any physical or metaphysical idea motivated by or founded upon the concept of incompatibility now requires careful re-examination. Three 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 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 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,

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

which merely expresses an idea most physicists—including quantum theorists—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 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 $\mathscr{A} + \mathscr{B} \leftrightarrow S$. 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! Thus when P1S is replaced by P1, it becomes apparent that the search for 'quantum logics' receives no legitimate motivation or endorsement from quantum physics.

(3) The 'microcausality' principle (Mandl, 1959; Streater & Wightman, 1964) 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 superficial argument involved is nothing but a pun on the term 'interference'. Consider two local observables $\mathscr{A}(x_1)$, $\mathscr{B}(x_2)$, i.e., observables associated with space-time points x_1, x_2. According to the theory of relativity, no signal can connect x_1 and x_2 if they are separated by a space-like interval in Minkowski space. Hence no operation at x_1 could possibly 'interfere' with an operation at x_2. On the other hand, if $[A(x_1),B(x_2)]\neq 0$, then the orthodox version of quantum theory insists that $\mathscr{A}(x_1)$ and $\mathscr{B}(x_2)$ 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 x_1 and x_2 have a space-like separation, relativity precludes any 'interference' between the measurement operations; hence $[A(x_1),B(x_2)]=0$ 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.

Much of the argument here criticized would hold if the term measurement were replaced by preparation of state.

8. *Joint Probability in Quantum Theory*

Research concerning joint probability distributions of noncommuting quantum observables has been reported by various authors

(Dirac, 1945; Moyal, 1949; Suppes, 1949, 1966; Wigner, 1932; Margenau & Hill, 1961; Margenau, 1963; Cohen, 1966; Cohen and Margenau, 1967). Interestingly, their motivations seem to rest upon rather diverse problems—hidden variables, operator-observable correspondence, reduction of diagonal matrix elements to phase space integrals, etc. However, the common feature of all these attempts to derive joint distributions has been their failure to meet all the mathematical conditions to be imposed on distribution functions. 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. A natural starting point for such an investigation is the consistency condition (c_1):

If $\mathscr{F}(\mathscr{A},\mathscr{B}) \leftrightarrow F$ and $W(a_k,b_l;\rho)$ is the (unknown) joint probability that $\mathscr{M}_1(\mathscr{A})$, $\mathscr{M}_1(\mathscr{B})$ would yield a_k, b_l, respectively, then

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

It is also natural to assume, as we have already done implicitly in previous sections, that the marginal distribution associated with $W(a_k,b_l;\rho)$ should equal the quantum mechanical distributions associated with single measurements; i.e.,

$$\sum W(a_k,b_l;\rho) = \mathrm{Tr}\,(\rho P_{\alpha_k})$$

$$\sum_{k} W(a_k,b_l;\rho) = \mathrm{Tr}\,(\rho P_{\beta_l})$$

Strictly speaking these equations are special cases of (c_1).

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 $\mathscr{A}$-measurement results might depend on whether or not the $\mathscr{A}$-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 $\mathscr{M}_1(\mathscr{A})$ yields elements of $\mathscr{E}(A)$, *provided* no physical operation except the $\mathscr{A}$-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 $\mathscr{A}$-measurement 'by itself' and an $\mathscr{A}$-measurement 'in conjunction with other operations'. Accordingly, for the sake of simplicity we assume, as is always done, that 'an $\mathscr{A}$-measurement is an $\mathscr{A}$-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 $\mathscr{A},\mathscr{B}$-measurements upon the members of a single ensemble (compare 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 $\mathscr{A} = \mathscr{X}$ (position) and $\mathscr{B} = \mathscr{P}$ (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

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp w(x,p;\psi)\mathscr{F}(x,p) = \langle \psi, F\psi \rangle$$

Now, *if* there were a general rule which provided F given $\mathscr{F}(x,p)$ 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* $\mathscr{F} \leftrightarrow F$ *turned out to be unsatisfactory.* Previous publications show this to be the case.

Suppes (1961) has argued along these lines that quantum mechanics cannot admit a proper distribution $w(x,p;\psi)$ and that simultaneous measurement of $\mathscr{X}$ and $\mathscr{P}$ 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

$$\mathscr{F}(x,p) = \exp(i\theta x + i\tau p)$$

Suppes assumes without special comment that if $\mathscr{X} \leftrightarrow X, \mathscr{P} \leftrightarrow P$, then

$$\mathscr{F}(\mathscr{X},\mathscr{P}) = \exp(i\theta\mathscr{X} + i\tau\mathscr{P}) \leftrightarrow \exp(i\theta X + i\tau P) = F$$

This suffices to determine $w(x,p;\psi)$ uniquely as the Fourier transform of $\langle\psi,\exp(i\theta X+i\tau P)\psi\rangle$:

$$w(x,p;\psi)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\psi^*\left(x-\frac{u\hbar}{2}\right)\exp(-iup)\,\psi\left(x+\frac{u\hbar}{2}\right)du$$

Because this w is not positive semidefinite† yet is unique, Suppes concludes that $\mathscr{X}$ and $\mathscr{P}$ are incompatible observables.

Having seen the same erroneous conclusion drawn from (c_1) before (compare Section 4), it is not so difficult this time to reinterpret the mathematics. Suppes has not proved the sweeping assertion that $\mathscr{X}$ and $\mathscr{P}$ cannot be measured simultaneously; he has shown only that the correspondence

$$\exp(i\theta\mathscr{X}+i\tau\mathscr{P})\leftrightarrow\exp(i\theta X+i\tau P)$$

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 (1966) deduces from several premises that 'the functional or working logic of quantum mechanics is not classical'. Among these premises is his (false) assumption, motivated as above, 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 and Hill (1961), begins with the concept of covariance:

$$\begin{aligned}\mathrm{Cov}(\mathscr{A},\mathscr{B})&=\langle\mathscr{A}\mathscr{B}\rangle-\langle\mathscr{A}\rangle\langle\mathscr{B}\rangle\\&=\sum_{kl}W(a_k,b_l;\psi)\,a_k b_l-\langle\mathscr{A}\rangle\langle\mathscr{B}\rangle\end{aligned}$$

The first problem is to find an Hermitean operator $F\leftrightarrow\mathscr{A}\mathscr{B}$ which leads to desirable properties for $\mathrm{Cov}(\mathscr{A},\mathscr{B})$. Once F is selected, the two covariance expressions above are equated to obtain what will be recognized as the (c_1) condition for $\mathscr{F}(\mathscr{A},\mathscr{B})=\mathscr{A}\mathscr{B}$:

$$\sum_{kl}W(a_k,b_l;\psi)\,a_k b_l=\langle\psi,F\psi\rangle$$

Because it gives the covariance reasonable properties, Margenau and Hill chose

$$F=\tfrac{1}{2}(AB+BA)$$

† This w, known as the Wigner distribution (Wigner, 1932), is a useful computational tool in statistical mechanics. It is not the only function serving this purpose, as will be noted presently.

Unfortunately, the W's to which this leads may take on negative values, as in the other instances.

With our reinterpretation of von Neumann's 'simultaneous measurability' theorem (compare Section 7), perhaps it is possible to understand why this covariance method was unsuccessful. Since that theorem, correctly interpreted, proves that the correspondence $\mathscr{A}\mathscr{B} \leftrightarrow \frac{1}{2}(AB + BA)$ 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 conclusions of Margenau and Hill may 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. This was already done in earlier investigations (Margenau & Hill, 1961; Margenau, 1963). 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 namely $W'(a_k, b_l; \psi) = |\langle \alpha_k, \psi| \rangle^2 |\langle \beta_l, \psi \rangle|^2$. As shown by Margenau (1963), this stochastically independent form of W' in no way violates the spirit of quantum theory. (The uncertainty principle, for example, is derivable from W', a fact which shows again how improper 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' unique?

For the special case where $A = X, B = P, [X,P] = i\hbar 1$, Cohen (1966) has laid the ground-work for answering this question by providing a canonical form for $w(x,p;\psi)$:

$$w(x,p;\psi) = \frac{1}{4\pi^2} \int d\theta \int du \int d\tau \exp(-i\theta x - i\tau p + i\theta u)$$

$$\times f(\theta,\tau)\, \psi^*\left(u - \frac{\tau\hbar}{2}\right) \psi\left(u + \frac{\tau\hbar}{2}\right)$$

any choice of $f(\theta,\tau)$ such that $f(0,0)=f(\theta,\tau)=1$ will generate a w satisfying the marginal requirements. The f for any given w is found by Fourier inversion. No way has been found to isolate those f's which generate positive semidefinite w's.

Nevertheless there *exist* semidefinite joint probability distributions for every state ψ, as we shall indicate by way of examples. Apparently they cannot be specified by the axioms of quantum mechanics, nor by any elegant general rule we have discovered. The reason for this, it seems, is the fact that compound operations do have results which often depend upon physical circumstances, on the nature of states as well as the manner of observation. Hence one would not expect a general set of axioms to control their mathematical representations. In an earlier paper (Margenau, 1963), note was taken of these vagaries concerning joint measurements by introducing the concept of a *compound* measurement, which avoids the need for defining joint probabilities altogether. Our present pursuit is parallel, but along more traditional lines.

If we restrict ourselves to operators with discrete spectra, the problem of finding *all* admissible $W(a_k,b_l;\psi)$ becomes more tractable. Consider the special case $\mathscr{A}=(2/\hbar)\mathscr{S}_z$, $\mathscr{B}=(2/\hbar)\mathscr{S}_x$, where $\mathscr{S}_x$, $\mathscr{S}_z$ are spin components for a spin-$\frac{1}{2}$ system. The operators which correspond to $\mathscr{A}$ and $\mathscr{B}$ are represented by the Pauli matrices:

$$\mathscr{A}\leftrightarrow\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}=\boldsymbol{A}, \qquad \mathscr{B}\leftrightarrow\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}=\boldsymbol{B}$$

If

$$\alpha_1=\begin{pmatrix}1\\ 0\end{pmatrix}, \qquad \alpha_2=\begin{pmatrix}0\\ 1\end{pmatrix}, \qquad \beta_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}, \qquad \beta_2=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix},$$

$$a_1=b_1=1, \qquad a_2=b_2=-1$$

these eigenvalue equations hold:

$$A\alpha_k=a_k\alpha_k, \qquad B\beta_l=b_l\beta_l$$

We seek all functions $W(a_k,b_l;\psi)\equiv W_{kl}(\psi)$ satisfying the following conditions:

$$0\leqslant W_{kl}(\psi)\leqslant 1, \qquad \sum_l W_{kl}(\psi)=|\langle\alpha_k,\psi\rangle|^2, \qquad \sum_k W_{kl}(\psi)=|\langle\beta_l,\psi\rangle|^2$$

For the present simple case, it is in fact possible to find all W's consistent with any given ψ. Consider, for example,

$$\psi=\begin{pmatrix}\sqrt{0{\cdot}9}\\ \sqrt{0{\cdot}1}\end{pmatrix}$$

The marginal conditions are then

$$\begin{aligned} W_{11} + W_{12} \qquad\qquad &= 0{\cdot}9 \\ W_{21} + W_{22} &= 0{\cdot}1 \\ W_{11} \qquad + W_{21} \qquad &= 0{\cdot}8 \\ W_{12} \qquad + W_{22} &= 0{\cdot}2 \end{aligned}$$

Substituting $W_{22} = 0{\cdot}2 - W_{12}$ into the second equation, we obtain three equations in three unknowns W_{11}, W_{12}, W_{21}:

$$\begin{aligned} W_{11} + W_{12} \qquad\qquad &= 0{\cdot}9 \\ - W_{12} + W_{21} &= -0{\cdot}1 \\ W_{11} \qquad + W_{21} &= 0{\cdot}8 \end{aligned}$$

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 $0 \leqslant W_{kl} \leqslant 1$ is used to delineate the segment of that line which represents *all* admissible sets (W_{11}, W_{12}, W_{21}) and W_{22} (via $W_{22} = 0{\cdot}2 - W_{12}$). When this is done, there results a one parameter (τ) family of positive semidefinite W's satisfying the marginal conditions:

$$W_\tau(a_k, b_l; \psi) = m_{kl}\tau + n_{kl}, \qquad \tau \in [0, 1]$$

$$(m_{kl}) = \begin{pmatrix} 0{\cdot}1 & -0{\cdot}1 \\ -0{\cdot}1 & 0{\cdot}1 \end{pmatrix}, \qquad (n_{kl}) = \begin{pmatrix} 0{\cdot}7 & 0{\cdot}2 \\ 0{\cdot}1 & 0 \end{pmatrix}$$

(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 space.)

One member of this family is the uncorrelated distribution

$$W'(a_k, b_l; \psi) = |\langle \alpha_k, \psi \rangle|^2 \, |\langle \beta_l, \psi \rangle|^2$$

to find the corresponding parameter τ' consider the form

$$W_{\tau'}(a_k, b_l) = f_k g_l$$

From the equations,

$$\begin{aligned} f_1 g_1 &= 0{\cdot}1\tau + 0{\cdot}7, \qquad & f_2 g_1 &= -0{\cdot}1\tau + 0{\cdot}1 \\ f_1 g_2 &= -0{\cdot}1\tau + 0{\cdot}2, \qquad & f_2 g_2 &= 0{\cdot}1\tau \end{aligned}$$

two expressions for g_1/g_2 may be derived:

$$\frac{g_1}{g_2} = \frac{0{\cdot}1\tau + 0{\cdot}7}{-0{\cdot}1\tau + 0{\cdot}2} = \frac{-0{\cdot}1\tau + 0{\cdot}1}{0{\cdot}1\tau}$$

Solving for τ, we obtain $\tau' = 0{\cdot}2$. 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, $W(a_k, b_l; \psi)$. In general, given any state ρ, we therefore anticipate the existence of many distribution functions marginally consistent with that state. Again, we are reminded of the ingression of physical circumstances into our formal scheme.

Nevertheless, it is difficult to evaluate fully the significance of this apparent multiplicity of W's. There are in principle two possibilities: (1) For every state there exists a unique $W(a_k, b_l; \rho)$ governing the statistics of $\mathscr{A}$,$\mathscr{B}$-measurement results for real physical systems; or (2) The quantum mechanical state ρ does not determine the joint distribution of $\mathscr{A}$,$\mathscr{B}$-measurement results.

Alternative (1) has always been assumed implicitly throughout most 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 $W(a_k, b_l)$ just as it determines $W_{\mathscr{A}}(a_k)$ and $W_{\mathscr{B}}(b_l)$. 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

$$\mathscr{A} \leftrightarrow \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \qquad \mathscr{B} \leftrightarrow \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}, \; a_1 \neq a_2, \; b_1 \neq b_2$$

whose common eigenvectors are

$$\alpha_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

For the state

$$\psi = \begin{pmatrix} \sqrt{\frac{3}{4}} \\ \sqrt{\frac{1}{4}} \end{pmatrix}$$

we obtain these marginal distributions:

$$W_{11} + W_{12} = W_{11} + W_{21} = \tfrac{3}{4}$$
$$W_{21} + W_{22} = W_{12} + W_{22} = \tfrac{1}{4}$$

These equations are satisfied by the following one parameter family of W's:

$$W_{kl} = m_{kl}\tau + n_{kl}, \qquad \tau \in [0, 1]$$

where

$$(m_{kl}) = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix}, \qquad (n_{kl}) = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 \end{pmatrix}$$

Nevertheless, we are able to select from among these, by invoking further criteria, the single correct joint distribution. By working with an auxiliary observable $\mathscr{C}$ through which $\mathscr{A}$ and $\mathscr{B}$ may be jointly measured in the trivial sense, we find that

$$W_{11} = \tfrac{3}{4}, \qquad W_{22} = \tfrac{1}{4}, \qquad W_{12} = W_{21} = 0$$

I.e., $\tau = 1$ 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.) But this example hardly suggests a *general* procedure.

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 *not* 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 *as a functional of state*, that simultaneous measurements are in some sense theoretical anomalies not tractable by any *comprehensive*, universal 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 roam in the relatively unexplored (due to aforementioned taboos) realm of simultaneous measurement theory.

9. *A Search for 'Simple' 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 doctrine of physical observables of *latency* (Margenau, 1950). 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 ...'). 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 $\mathscr{M}_1$'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 introduced by one of the authors (Margenau, 1950)]. To signalize the logical distinction we designate the latter class of constructs, which form part of the *theoretical* structure of our problem, by $\mathscr{M}_2$. These constructs were exemplified in Section 5 by a discussion of the time-of-flight $\mathscr{M}_2(\mathscr{P})$ for momentum measurement; typically, the theory of $\mathscr{M}_2(\mathscr{P})$ established a connection between $\mathscr{M}_1(\mathscr{X})$ already operationally defined, and $\mathscr{M}_1(\mathscr{P})$, which thereby acquired a *theoretically validated* definition itself.

In previous discussions of the joint-measurement problem empirical examples, incompletely analyzed, have borne the brunt of the arguments in favor of the thesis we are here advancing. There *are* procedures, it was pointed out, which permit an assignment of values to pairs of noncommuting observables. Our present aim goes beyond such indications; it is to show within the context of measurement *theory*, as presented in the foregoing pages, how such empirical operations function as parts of 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 $\mathscr{M}_2$.

That is, if a given procedure $\mathscr{M}_2(\mathscr{X},\mathscr{P})$ is to be regarded as a method for simultaneous measurement of $\mathscr{X}$ and $\mathscr{P}$, the scheme must be certified by a theory establishing relations between $\mathscr{M}_1(\mathscr{X})$, $\mathscr{M}_1(\mathscr{P})$ and whatever 'direct meter readings' are used as the basis for inference of simultaneous $\mathscr{M}_1(\mathscr{X})$ and $\mathscr{M}_1(\mathscr{P})$-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 $\mathscr{X}$- and $\mathscr{P}$-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 $\mathscr{M}_2$-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 $\mathscr{M}_2$ 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 $\mathscr{M}_1$-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 belonging to type A. (2) An *historical* $\mathscr{M}_2$-theory also seeks to certify some operation as a *bona fide* supplier of numbers which can be meaningfully interpreted as $\mathscr{M}_1$-results for S in state ρ_{t_0}. However, unlike the simple type A, the historical $\mathscr{M}_2$-theory cannot be worked 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 simple time-of-flight $\mathscr{M}_2(\mathscr{P})$ and the historical time-of-flight $\mathscr{M}_2(\mathscr{X},\mathscr{P})$ of type B.

Physically, the $\mathscr{M}_2$-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. 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 $\mathscr{M}_1$'s usually presupposes that measurements are performed upon systems in states which are simply given without

† I.e., practically arbitrary. Recall, e.g., that the time-of-flight $\mathscr{M}_2(\mathscr{P})$ did require $\psi(x)$ to have compact support, but no specific functional form was demanded.

details as to the actual method of preparation. Accordingly, $\mathscr{M}_2$-schemes for single observables (or commuting sets of observables) have been of type A. It is therefore natural to seek a simple $\mathscr{M}_2$-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 $\mathscr{M}_2(\mathscr{X},\mathscr{P})$ and the use of two systems already correlated at the time of interest—were of type B, there is perhaps no reason to be optimistic about finding any simple theory for simultaneous measurement.

We shall now outline two fairly general procedures which, at the outset, seem to be altogether plausible methods for achieving simultaneous type A measurements of two noncommuting observables. In both cases theoretical obstacles will eventually arise, and this will be interpreted as evidence that quantum theory does probably forbid simultaneous type A measurements. Deeper reasons to anticipate such a theoretic prohibition will then be explored.

Let S be the system upon which $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ is to be performed. Suppose a second system M with an observable $\mathscr{U}$, $U\theta_l = u_l\theta_l$ interacts with S in such a way that the following correlations result:

$$T_{\mathscr{A},\mathscr{B}}(\psi \otimes \chi_0) = \sum_k c_{kl}(\psi)\,\alpha_k \otimes \theta_l$$

with marginal conditions

$$\sum_l |c_{kl}(\psi)|^2 = |\langle \alpha_k, \psi\rangle|^2 \qquad \text{and} \qquad \sum_k |c_{kl}(\psi)|^2 = |\langle \beta_l, \psi\rangle|^2$$

where $T_{\mathscr{A},\mathscr{B}}$ is the evolution operator for this $\mathscr{M}_2(\mathscr{A},\mathscr{B})$-scheme and ψ and $\mathscr{X}_0$ are the initial state vectors of S and M. Since these equations imply that post-interaction $\mathscr{M}_1(\mathscr{A})$-results occur with the same probability as they would have before interaction and that a post-interaction $\mathscr{M}_1(\mathscr{U})$ would yield u_n with the same probability that a pre-interaction $\mathscr{M}_1(\mathscr{B})$ would have yielded b_n, a simple $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ procedure yielding $\mathscr{M}_1(\mathscr{A})$- and $\mathscr{M}_1(\mathscr{B})$-results for the instant before interaction may be defined as follows: after interaction, measure $\mathscr{A}$ on S, $\mathscr{U}$ on M; if the resulting number pair is (a_k,u_n), the pair (a_k,b_n) is declared the result of simultaneous $\mathscr{M}_1(\mathscr{A})$ and $\mathscr{M}_1(\mathscr{B})$ for S in the state ψ, i.e., just before interaction with M. If a $T_{\mathscr{A},\mathscr{B}}$ exists which can accomplish these correlations, the derivation of the associated joint probability distribution is immediate:

$$W(a_k,b_l;\psi;T_{\mathscr{A},\mathscr{B}}) = |c_{kl}(\psi)|^2$$

We have indicated a possible dependence of W upon $T_{\mathscr{A},\mathscr{B}}$, 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 $T_{\mathscr{A},\mathscr{B}}$ exists, we have an example of a type A theory of $\mathscr{M}_2(\mathscr{A},\mathscr{B})$, complete with joint probability distribution. Unfortunately, no such $T_{a,\mathscr{B}}$ exists.

To see why, consider first $\psi = \alpha_n$. In this case

$$\sum_l |c_{kl}(\alpha_n)|^2 = |\langle \alpha_k, \alpha_n \rangle|^2 = \delta_{kn}$$

Since $|c_{kl}|^2 > 0$, this implies that $c_{kl}(\alpha_n) = 0$, $k \neq n$; thus $c_{kl}(\alpha_n)$ has the form $c_{kl}(\alpha_n) = \delta_{kn} g_l(\alpha_n)$ and accordingly

$$\sum_k |c_{kl}(\alpha_n)|^2 = |g_l(\alpha_n)|^2 = |\langle \beta_l, \alpha_n \rangle|^2$$

Hence

$$T_{\mathscr{A},\mathscr{B}}(\alpha_n \otimes \chi_0) = \sum_{kl} \delta_{kn} g_l(\alpha_n)\, \alpha_k \otimes \theta_l$$

Now, by superposition we can derive the transformation associated with any ψ subjected to this type of measurement. If

$$\psi = \sum_n \langle \alpha_n, \psi \rangle \alpha_n$$

$$T_{\mathscr{A},\mathscr{B}}(\psi \otimes \chi_0) = \sum_{nkl} \langle \alpha_n, \psi \rangle \delta_{kn} g_l(\alpha_n)\, \alpha_k \otimes \theta_l$$

$$= \sum_{kl} \langle \alpha_k, \psi \rangle g_l(\alpha_k)\, \alpha_k \otimes \theta_l$$

i.e.,

$$c_{kl}(\psi) = g_l(\alpha_k) \langle \alpha_k, \psi \rangle$$

Checking this form against the two required marginal conditions, we find that one of them is not in general satisfied; to be sure,

$$\sum_l |c_{kl}(\psi)|^2 = \left(\sum_l |g_l(\alpha_k)|^2 \right) |\langle \alpha_k, \psi \rangle|^2 = |\langle \alpha_k, \psi \rangle|^2$$

but

$$\sum_k |c_{kl}(\psi)|^2 = |g_l(\alpha_k)|^2 \left(\sum_k |\langle \alpha_k, \psi \rangle|^2 \right) = |g_l(\alpha_k)|^2 = |\langle \beta_l, \alpha_k \rangle|^2 \neq |\langle \beta_l, \psi \rangle|^2$$

Hence the simple $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ here envisaged is theoretically impossible.

The same conclusion is obtained if we attempt to construct a type A $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ based upon interaction of the primary system S with *two* measuring devices, $M(\mathscr{A})$ and $M(\mathscr{B})$, which we might call $\mathscr{A}$- and $\mathscr{B}$-'meters'. If χ_0, ξ_0 are the initial states for the meters, $\mathscr{U}$, $\mathscr{V}$ are

observables belonging to $M(\mathscr{A})$, $M(\mathscr{B})$, respectively, $U\theta_m = u_m\theta_m$, $V\eta_n = v_n\eta_n$, and $\{\varphi_l\}$ is an orthonormal set of vectors in the Hilbert space of S, the following correlation scheme would describe an $\mathscr{M}_2(\mathscr{A},\mathscr{B})$:

$$T_{\mathscr{A},\mathscr{B}}(\psi \otimes \chi_0 \otimes \xi_0) = \sum_{lmn} s_{lmn}(\psi)\,\varphi_l \otimes \theta_m \otimes \eta_n$$

with marginal conditions

$$\sum_{ln} |s_{lmn}(\psi)|^2 = |\langle \alpha_m, \psi\rangle|^2$$

and

$$\sum_{lm} |s_{lmn}(\psi)|^2 = |\langle \beta_n, \psi\rangle|^2$$

$\mathscr{M}_2(\mathscr{A},\mathscr{B})$ would consist simply of 'reading' the two meters (measuring $\mathscr{U}$ and $\mathscr{V}$) after the interaction; if (u_m, v_n) results, (a_m, b_n) is declared the result of simultaneous $\mathscr{M}_1(\mathscr{A})$ and $\mathscr{M}_1(\mathscr{B})$ upon S in state ψ, i.e., just before the interaction. Once again the joint distribution would be easily derived:

$$W(a_m, b_n; \psi; T_{\mathscr{A},\mathscr{B}}) = \sum_{l} |s_{lmn}(\psi)|^2$$

Like the previous example, this $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ is impossible; thus noncommuting observables cannot be simultaneously measured merely by letting S interact with two meters and reading the two separate results. To prove this, note first that for $\psi = \alpha_k$,

$$T_{\mathscr{A},\mathscr{B}}(\alpha_k \otimes \chi_0 \otimes \xi_0) = \sum_{lmn} \delta_{mk}\, g_{ln}(\alpha_k)\,\varphi_l \otimes \theta_m \otimes \eta_n$$

$$= \sum_{ln} g_{ln}(\alpha_k)\,\varphi_l \otimes \theta_k \otimes \eta_n$$

$$\sum_{lm} |\delta_{mk}\, g_{ln}(\alpha_k)|^2 = \sum_{l} |g_{ln}(\alpha_k)|^2 = |\langle \beta_n, \alpha_k\rangle|^2$$

$$\sum_{ln} |\delta_{mk}\, g_{ln}(\alpha_k)|^2 = |\langle \alpha_m, \alpha_k\rangle|^2 = \delta_{mk}$$

Now consider the superposition

$$\psi = \sum_k \langle \alpha_k, \psi\rangle\, \alpha_k$$

It leads to

$$T_{\mathscr{A},\mathscr{B}}(\psi \otimes \chi_0 \otimes \xi_0) = \sum_{kln} \langle \alpha_k, \psi\rangle\, g_{ln}(\alpha_k)\,\varphi_l \otimes \theta_k \otimes \eta_n$$

i.e.,

$$s_{lkn} = g_{ln}(\alpha_k)\langle \alpha_k, \psi\rangle$$

But this s_{lkn} does not satisfy the second marginal condition:

$$\sum_{lk} |s_{lkn}|^2 = \sum_{lk} |\langle \alpha_k, \psi \rangle|^2 |g_{ln}(\alpha_k)|^2 = \sum_{k} |\langle \alpha_k, \psi \rangle|^2 \sum_{l} |g_{ln}(\alpha_k)|^2$$
$$= \sum_{k} |\langle \alpha_k, \psi \rangle|^2 |\langle \beta_n, \alpha_k \rangle|^2 \neq \left| \sum_{k} \langle \alpha_k, \psi \rangle \langle \beta_n, \alpha_k \rangle \right|^2 = |\langle \beta_n, \psi \rangle|^2$$

The foregoing examples were presented as evidence for this proposition: (O) Simultaneous type A measurements of noncommuting observables are theoretically impossible.

Of course merely citing two unsuccessful attempts to develop a simple $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ does not prove the proposition, but for the first time in the present study, we have found 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 $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ 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 $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ which leads to numbers (α_k, b_l) interpretable as $\mathscr{M}_1(\mathscr{A})$- and $\mathscr{M}_1(\mathscr{B})$-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, simple $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ 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, $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ may be performed and will yield (a_k, b_l), the results of simultaneous $\mathscr{M}_1(\mathscr{A})$, $\mathscr{M}_1(\mathscr{B})$ 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 $\mathrm{Tr}_C\, \rho_{S+C}(t_0) = \rho_S(t_0)$, where Tr_C denotes the trace over C's Hilbert space.

Since S is not in interaction with C, the efficacy of $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ 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-Podolski-Rosen (Einstein *et al.*, 1935) (E.P.R.) discussion concerning the completeness of quantum mechanics. If $\mathscr{A}, \mathscr{B}$ are the S-observables to be measured and $\mathscr{U}$, $\mathscr{V}$ are noncommuting C-observables with respective eigenvector sets $\{\theta_m\}$, $\{\eta_n\}$, such double correlation may be expressed† as follows:

$$\rho_{S+C} = P_\Psi, \qquad \Psi = \sum_{k} c_k \alpha_k \otimes \theta_k = \sum_{l} d_l \beta_l \otimes \eta_l$$

† For specific examples, see Einstein *et al.* (1935) or Park (1968).

The first form of Ψ implies that from an $\mathscr{M}_1(\mathscr{A})$-result one may conclude what result a concurrent $\mathscr{M}_1(\mathscr{U})$ would have yielded; similarly the second form shows that if $\mathscr{M}_1(\mathscr{B})$ is performed the result that a simultaneous $\mathscr{M}_1(\mathscr{V})$ would yield may be inferred. Hence if the procedure $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ is applied to S, the results (a_k,b_l) lead to the *prediction with certainty* that simultaneous $\mathscr{M}_1(\mathscr{U})$ and $\mathscr{M}_1(\mathscr{V})$ would have yielded (u_k,v_l) at t_0. But system C is not disturbed by this process; during $\mathscr{M}_2(\mathscr{A},\mathscr{B})$, C evolves from its state at t_0 just as it would have if $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ had not been performed on S. However, $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ has provided a means for dividing the C-ensemble into identifiable subensembles. In particular, consider the subensemble comprised of those members of the C-ensemble for which $\mathscr{M}_1(\mathscr{U})$ was certain to yield u_K and $\mathscr{M}_1(\mathscr{V})$ 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:

$$W_{\mathscr{U}}[u_K;\rho_{KL}(t_0)] = \mathrm{Tr}\,[\rho_{KL}(t_0)P_{\theta_K}] = 1$$

$$W_{\mathscr{V}}[v_L;\rho_{KL}(t_0)] = \mathrm{Tr}\,[\rho_{KL}(t_0)P_{\eta_L}] = 1$$

and by Th3, $\mathrm{Tr}\,[\rho_{KL}(t_0)] = 1$. By using the matrix representation of these equations, it is easy to see that the only possible solution is

$$\rho_{KL}(t_0) = P_{\theta_K} = P_{\eta_L}$$

But since $[\mathscr{U},\mathscr{V}] \neq 0$, 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 $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ performable upon any arbitrary $\rho_S(t_0)$, i.e., an $\mathscr{M}_2(\mathscr{A},\mathscr{B})$ of type A; (b) for every $\rho_S(t_0)$, it is possible to find $\rho_{S+C}(t_0)$ which incorporates the EPR-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 $\mathrm{Tr}_C P_\Psi = P_\psi$. 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 $\mathscr{A}$- and $\mathscr{U}$-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 (O) conclusively.

It does, however, prove this: (O′) *if* a system is in any doubly correlated state $\rho_S(t_0)$, then a simultaneous type A 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 $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ performable upon any $\rho_S(t_0)$ *except those which are doubly correlated in* $\mathscr{A}$ *and* $\mathscr{B}$—a rather strange statement.

Although the foregoing considerations do not rigorously prove that type A $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$ are generally impossible, it is difficult to allay the suspicion that (O) may indeed be correct. The above reported failure of reasonable attempts to develop *simple* $\mathscr{M}_2$-theories together with the proof just given of (O′) certainly point in the direction of (O). Moreover, the fact that we have been unable to find even specific cases of *simple* $\mathscr{M}_2(\mathscr{A}, \mathscr{B})$'s provides inductive support for (O).

Finally, it is instructive to summarize in a concise way the theoretical basis of our 'partial proof' of (O), i.e., our proof of (O′). 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 $\mathscr{X}$- and $\mathscr{P}$-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 (Section 1) that measurement and preparation must not be equated as is customary among 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 type A 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 type A 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. Because of its practical, and perhaps general, validity for type A 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 such simple joint $\mathscr{M}_2$'s were nonexistent, this would not eliminate *all* simultaneous measurement schemes. There would still be interesting type B procedures to study. Accordingly, the remainder of this section will be devoted to derivations of the joint distributions associated with several joint $\mathscr{M}_2$'s of this historical type. Perhaps another suitable adjective to describe this class of $\mathscr{M}_2$'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 $\mathscr{M}_2(\mathscr{X},\mathscr{P})$: The theory behind this method was presented in Section 5, where is was shown that, under certain conditions, an $\mathscr{X}$-measurement yielding x implied that a simultaneous $\mathscr{P}$-measurement would have yielded $p = mx/t$ 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 $\mathscr{M}_2(\mathscr{X},\mathscr{P})$ 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 $\mathscr{M}_2(\mathscr{X},\mathscr{P})$: Consider an electron (charge e) in a 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:

$$\psi(x,t) = \sqrt{\left(\frac{m}{2\pi i\hbar t}\right)}\exp\frac{i}{\hbar}\left[\frac{mx^2}{2t} + \tfrac{1}{2}ftx - \frac{f^2t^3}{24m}\right]$$

$$\times \int_{-\infty}^{\infty} dx_1 \exp\frac{i}{\hbar}\left[-\frac{mxx_1}{t} + \frac{mx_1{}^2}{2t} + \frac{ftx_1}{2}\right]\psi(x_1,0)$$

where $f = eE$.

From this transformation follow two probability 'matchings' on which an $\mathscr{M}_2(\mathscr{X},\mathscr{P})$ may be based:

(a) $W_{\mathscr{P}}[p \in (p_1,p_2);\psi_0] = W_{\mathscr{P}}[p \in (p_1+ft, p_2+ft;\psi_t)]$

(b) $W_{\mathscr{P}}[p \in (p_1,p_2);\psi_0] = W_{\mathscr{X}}\left[x \in \left(\frac{p_1t}{m} + \frac{ft^2}{2m}, \frac{p_2t}{m} + \frac{ft^2}{2m}\right);\psi_t\right]$,

as $t \to \infty$

provided $\psi(x,0)$ is nonzero only in some finite interval $(-x_0,x_0)$.

Equation (a) means that $\mathscr{M}_1(\mathscr{P})$ at $t=0$ yields p with the same probability that $\mathscr{M}_1(\mathscr{P})$ at $t>0$ yields $p+ft$. Equation (b) provides this $\mathscr{M}_2(\mathscr{P})$: to determine what $\mathscr{M}_1(\mathscr{P})$ at $t=0$ would have yielded, measure $\mathscr{X}$ at sufficiently large t and use the result x to evaluate $p = (mx/t) - (ft/2)$, which then counts as the result of $\mathscr{M}_1(\mathscr{P})$ at $t=0$. By analogy to the time-of-flight method, we have only to combine (a) and (b) to obtain an $\mathscr{M}_2(\mathscr{X},\mathscr{P})$: 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 $\mathscr{X}$; if the result is x, it is inferred through (a) and (b) that a simultaneous $\mathscr{M}_1(\mathscr{P})$ 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 $\psi(x,0)=(2\pi\hbar)^{-1}\exp(ipx/\hbar)$ and determine the corresponding $\psi(x,t)$:

$$\begin{aligned}\psi(x,t)&=\left\{\frac{1}{\sqrt{(2\pi\hbar)}}\sqrt{\left(\frac{m}{2\pi i\hbar t}\right)}\exp\frac{i}{\hbar}\left[\frac{mx^2}{2t}+\tfrac{1}{2}ftx-\frac{f^2t^3}{24m}\right]\right\}\\&\qquad\times\int_{-\infty}^{\infty}dx_1\exp\frac{i}{\hbar}\left[-\frac{mxx_1}{t}+\frac{m{x_1}^2}{2t}+\tfrac{1}{2}ftx_1+px_1\right]\\&=\left\{\ \right\}\int_{-\infty}^{\infty}dx_1\exp\left[-{x_1}^2\left(\frac{m}{2i\hbar t}\right)+ix_1\left(-\frac{mx}{t}+\frac{1}{2}\frac{ft}{\hbar}+\frac{p}{\hbar}\right)\right]\\&=\left\{\ \right\}\sqrt{\left(\frac{2\pi i\hbar t}{m}\right)}\exp\left[-\frac{it}{2m\hbar}\left(\tfrac{1}{2}ft-\frac{mx}{t}+p\right)^2\right]\\&=\left[\frac{1}{\sqrt{(2\pi\hbar)}}\exp\frac{i}{\hbar}(p+ft)x\right]\left[\exp-\frac{it}{\hbar}\left(\frac{p^2}{2m}+\frac{ftp}{2m}+\frac{f^2t^2}{6m}\right)\right]\end{aligned}$$

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:

$$\begin{aligned}&W_x\left[x\in\left(\frac{p_1t}{m}+\frac{ft^2}{2m},\frac{p_2t}{m}+\frac{ft^2}{2m}\right);\psi_t\right]\\&\quad=\int_{a=(p_1t/m)+(ft^2/2m)}^{b=(p_2t/m)+(ft^2/2m)}|\psi(x,t)|^2\,dx\\&\quad=\int_a^b\frac{m}{2\pi\hbar t}\left|\int_{-\infty}^{\infty}\exp\frac{i}{\hbar}\left[-\frac{mxx_1}{t}+\frac{m{x_1}^2}{2t}+\frac{ftx_1}{2}\right]\psi(x_1,0)\,dx_1\right|^2dx\end{aligned}$$

Under the substitution $x=(pt/m)+(ft^2/2m)$, this becomes

$$\int_{p_1}^{p_2}\left|\frac{1}{\sqrt{(2\pi\hbar)}}\int_{-\infty}^{\infty}\exp\left(-\frac{ipx_1}{\hbar}\right)\psi(x_1,0)\exp\left(\frac{im{x_1}^2}{2\hbar t}\right)dx_1\right|^2dp$$

which, by the same argument given in Section 5 in the time-of-flight case, for sufficiently large t becomes arbitrarily close to

$$\int_{p_1}^{p_2} \left| \frac{1}{\sqrt{(2\pi\hbar)}} \int_{-\infty}^{\infty} \exp\left(-\frac{ipx_1}{\hbar}\right) \psi(x_1, 0)\, dx_1 \right|^2 dp = W_{\mathscr{P}}[p \in (p_1, p_2); \psi_0]$$

(3) Correlated systems $\mathscr{M}_2(\mathscr{A}_1, \mathscr{B}_1)$: 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

$$\rho_1 = \mathrm{Tr}_2 P_\Psi = \sum_k |c_k|^2 P_{\alpha_k^{(1)}}$$

Now, the $\mathscr{M}_1(\mathscr{A}_2)$ which determines the desired $\mathscr{M}_1(\mathscr{A}_1)$-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 $\mathscr{M}_1(\mathscr{A}_1)$-result $a_k^{(1)}$ is assigned has state $P_{\alpha_k^{(1)}}$, and $\mathscr{M}_1(\mathscr{B}_1)$ on that subensemble would yield $b_l^{(1)}$ with probability $\mathrm{Tr}\,(P_{\alpha_k^{(1)}} P_{\beta_l^{(1)}}) = |\langle \beta_l^{(1)}, \alpha_k^{(1)} \rangle|^2$. 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 $\mathscr{M}_2$ is

$$W(a_k^{(1)}, b_l^{(1)}; \rho_1) = |c_k|^2 |\langle \beta_l^{(1)}, \alpha_k^{(1)} \rangle|^2 = \mathrm{Tr}\,(\rho_1 P_{\alpha_k^{(1)}})\, \mathrm{Tr}\,(P_{\alpha_k^{(1)}} P_{\beta_l^{(1)}})$$

To summarize, the foregoing joint measurement schemes led to these probability functions:

(1) $w(x, p; \psi_t) = |\psi_t(x)|^2 \delta\left(p - \dfrac{mx}{t}\right), \qquad t \to \infty$

(2) $w(x, p; \psi_t) = |\psi_t(x)|^2 \delta\left(p - \dfrac{mx}{t} - \dfrac{ft}{2}\right), \qquad t \to \infty$

(3) $W(a_k^{(1)}, b_l^{(1)}; \rho_1) = \mathrm{Tr}\,(\rho_1 P_{\alpha_k^{(1)}})\, \mathrm{Tr}\,(P_{\alpha_k^{(1)}} P_{\beta_l^{(1)}})$

Note that in each case the two observables are stochastically dependent.

(4) Magnetic field $\mathscr{M}_2(\mathscr{Y}, \mathscr{V}_y)$: 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*, $\mathscr{Y}$ and

$\mathscr{V}_y$. If an electron in a uniform magnetic field $\mathbf{B} = B\hat{k}$ is initially in the state $\Psi(x,y,z;0)$, the state at time t is given by (Einstein *et al.*, 1935)

$$\Psi(x,y,z;t) = \left(\frac{m}{2\pi i\hbar t}\right)^{3/2}\left(\frac{\omega t/2}{\sin \omega t/2}\right) \times \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dy_1 \int_{-\infty}^{\infty} dz_1 \exp\frac{im}{2\hbar}\left\{\frac{(z-z_1)^2}{t} + \frac{\omega}{2}\left(\cot\frac{\omega t}{2}\right) \times [(x-x_1)^2 + (y-y_1)^2] + \omega(x_1 y - x y_1)\right\} \Psi(x,y,z;0)$$

where $\omega \equiv eB/mc$ and the gauge is so chosen that the vector potential is $\mathbf{A} = B/2(-y\hat{\imath} + x\hat{\jmath})$. From this transformation the following probability equations may be derived:

(a) $W_{\mathscr{V}_y}[v_y \in (v_1, v_2); \Psi_0] = W_{\mathscr{V}_y}[v_y \in (-v_2, -v_1); \Psi_\tau]$

where $\tau \equiv \pi/\omega$.

(b) If

$$\Psi(x,y,z;0) = \delta(x)\psi(y)\zeta(z)$$

then

$$W_{\mathscr{V}_y}[v_y \in (v_1, v_2); \Psi_0] = W_{\mathscr{X}}\left[x \in \left(\frac{2v_1}{\omega}, \frac{2v_2}{\omega}\right); \Psi_\tau\right]$$

Proceeding as in previous examples, we may define an $\mathscr{M}_2(\mathscr{Y}, \mathscr{V}_y)$ as follows: to measure $\mathscr{Y}$ and $\mathscr{V}_y$ at time τ for an electron initially in the state $\delta(x)\psi(y)\zeta(z)$, measure $\mathscr{X}$ at τ and use the result x to evaluate $v_y = -(\omega x/2)$; also measure $\mathscr{Y}$ 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 $\mathscr{M}_1(\mathscr{Y})$ and $\mathscr{M}_1(\mathscr{V}_y)$ 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

$$\Psi(x,y,z;0) = \delta(x)\psi(y)\zeta(z)\dagger$$

It is convenient to work with the expansion

$$\Psi(x,y,z;0) = \delta(x)\int_{-\infty}^{\infty} \varphi(p)\frac{\exp(ipy/\hbar)}{\sqrt{(2\pi\hbar)}}dp\,\zeta(z)$$

† $\psi(y)$ and $\zeta(z)$ 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 =.)

where

$$\varphi(p) = \frac{1}{\sqrt{(2\pi\hbar)}} \int_{-\infty}^{\infty} \exp\left(-\frac{ipy}{\hbar}\right) \psi(y)\, dy$$

is the canonical momentum space wave function corresponding to $\psi(y)$.

$$\Psi\left(x, y, z; \frac{\pi}{\omega}\right)$$

$$= \left(\frac{m\omega}{2\pi^2 i\hbar}\right)\left(\frac{\pi}{2}\right)\left[\sqrt{\left(\frac{m\omega}{2\pi^2 i\hbar}\right)} \int_{-\infty}^{\infty} \exp\frac{imw}{2\pi\hbar}(z - z_1)^2\, \zeta(z_1)\, dz_1\right]$$

$$\times \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dy_1 \exp\frac{im}{2\hbar}[\omega(x_1 y - x y_1)]\, \delta(x_1)$$

$$\times \int_{-\infty}^{\infty} dp\, \varphi(p) \frac{\exp(-ipy_1/\hbar)}{\sqrt{(2\pi\hbar)}}$$

$$= \left(\frac{m\omega}{4\pi i\hbar}\right)[\zeta(z, \tau)] \int_{-\infty}^{\infty} dp\, \frac{\varphi(p)}{\sqrt{(2\pi\hbar)}} \int_{-\infty}^{\infty} dy_1 \exp -\frac{i}{\hbar}\left[\frac{m\omega x}{2} + p\right] y_1$$

$$= \left(\frac{m\omega}{4\pi i\hbar}\right) \zeta(z, \tau) \int_{-\infty}^{\infty} dp\, \frac{\varphi(p)}{\sqrt{(2\pi\hbar)}}\, 2\pi\delta\left[\frac{1}{\hbar}\left(\frac{m\omega x}{2} + p\right)\right]$$

$$= \left[\sqrt{\left(\frac{m\omega}{2}\right)} \varphi\left(\frac{m\omega x}{2}\right)\right]\left(\sqrt{\left[\frac{m\omega}{4\pi\hbar}\right]}\right) \zeta(z, \tau)$$

(The factor in brackets is normalized.)

From this expression we next obtain

$$W_{\mathscr{X}}\left[x \in \left(\frac{2v_1}{\omega}, \frac{2v_2}{\omega}\right); \Psi_\tau\right]$$

$$= \int_{2v_1/\omega}^{2v_2/\omega} \left(\frac{m\omega}{2}\right)\left|\varphi\left(\frac{m\omega x}{2}\right)\right|^2 dx \left[\int_{-\infty}^{\infty} \frac{m\omega}{4\pi\hbar}\, dy\right]\left(\int_{-\infty}^{\infty} |\zeta(z, \tau)|^2\, dz\right)$$

$$= \int_{v_1}^{v_2} |\varphi(mv)|^2\, m\, dv\, [\delta(0)](1)$$

$$\text{`}=\text{'} \int_{v_1}^{v_2} |\varphi(mv)|^2\, m\, dv$$

which must be compared to $W_{\mathscr{P}_y}[v \in (v_1, v_2); \Psi_0]$, where

$$\Psi(x, y, z; 0) = \delta(x)\, \psi(y)\, \zeta(z)$$

The operator V_y such that $V_y \leftrightarrow \mathscr{V}_y$ is

$$V_y = \frac{1}{m} P_y - \frac{e}{mc} A_y = \frac{\hbar}{im} \frac{\partial}{\partial y} - \frac{\omega}{2} x$$

its eigenfunctions are

$$\eta_{x_0 p_0 z_0} = \delta(x - x_0) \frac{\exp\left[(i/\hbar) y p_{y_0}\right]}{\sqrt{(2\pi\hbar)}} \delta(z - z_0)$$

The V_y-eigenvalue corresponding to $\eta_{x_0 p_{y_0} z_0}$ is

$$\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0$$

$W_{\mathscr{V}_y}$ is then calculated as follows:

$$W_{\mathscr{V}_0}[v \in (v_1, v_2); \Psi_0]$$

$$= \int_{v_1}^{v_2} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 |\langle \eta_{x_0 p_{y_0} z_0}, \Psi_0 \rangle|^2 \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega x_0}{2}\right)\right]$$

$$= \int_{v_1}^{v_2} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \left| \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz \right.$$

$$\left. \times \delta(x - x_0) \frac{\exp\left[-(iy/\hbar) p_{y_0}\right]}{\sqrt{(2\pi\hbar)}} \delta(z - z_0) \delta(x) \psi(y) \zeta(z) \right|^2$$

$$\times \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right]$$

$$= \int_{v_1}^{v_2} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega x_0}{2}\right)\right]$$

$$\times \left| \delta(x_0) \frac{1}{\sqrt{(2\pi\hbar)}} \int_{-\infty}^{\infty} \exp\left(-\frac{i p_{y_0} y}{\hbar}\right) \psi(y)\, dy \right|^2$$

$$= \int_{v_1}^{v_2} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right] \delta(x_0) \delta(x_0) |\varphi(p_{y_0})|^2$$

$$= \int_{v_1}^{v_2} dv \int_{-\infty}^{\infty} dp_{y_0} \delta\left[v - \frac{p_{y_0}}{m}\right] \delta(0) \, |\varphi(p_{y_0})|^2$$

$$= \int_{v_1}^{v_2} |\varphi(mv)|^2 \, m \, dv$$

Comparison with the expression for

$$W_{\mathscr{X}}\left[x \in \left(\frac{2v_1}{\omega}, \frac{2v_2}{\omega}; \Psi_\tau\right)\right]$$

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 $\mathscr{M}_2$ under consideration.

$$W_{\mathscr{Y}_y}[v_y \in (-v_2, -v_1); \Psi_\tau]$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \, |\langle \eta_{x_0 p_{y_0} z_0}, \Psi_\tau \rangle|^2 \, \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{w x_0}{2}\right)\right]$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0 \, \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{w}{2} x_0\right)\right]$$

$$\times \left| \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz \, \delta(x - x_0) \frac{\exp(-i p_{y_0} y/\hbar)}{\sqrt{(2\pi\hbar)}} \right.$$

$$\left. \times \, \delta(z - z_0) \sqrt{\left(\frac{m\omega}{2}\right)} \varphi\left(\frac{m\omega x}{2}\right) \sqrt{\left(\frac{m\omega}{4\pi\hbar}\right)} \zeta(z, \tau) \right|^2$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \left(\frac{m\omega}{2}\right) \delta(p_{y_0}) \, \delta(p_{y_0})$$

$$\times \left| \varphi\left(\frac{m\omega x_0}{2}\right) \right|^2 \delta\left[v - \left(\frac{p_{y_0}}{m} - \frac{\omega}{2} x_0\right)\right]$$

$$= \int_{-v_2}^{-v_1} dv \int_{-\infty}^{\infty} dx_0 \frac{m\omega}{2} \left| \varphi\left(\frac{m\omega x_0}{2}\right) \right|^2 \delta(0) \, \delta\left[v + \frac{\omega}{2} x_0\right]$$

$$= \int_{-v_2}^{-v_1} |\varphi(-m\omega)|^2 \, m \, dv \, \delta(0)$$

$$\text{`}=\text{'} \int_{-v_2}^{-v_1} |\varphi(-mv)|^2 \, m \, dv$$

$$= \int_{v_1}^{v_2} |\varphi(mv)|^2 \, m \, dv$$

This equals $W_{\mathscr{Y}_y}[v_y \in (v_1, v_2); \Psi_0]$; hence (a) is valid.

The joint probability density $w(y, v_y; \Psi_\tau)$ for this $\mathscr{M}_2(\mathscr{Y}, \mathscr{V}_y)$ is easily obtained. Since $\mathscr{M}_1(\mathscr{V}_y)$-results depend on $\mathscr{M}_1(\mathscr{X})$-results,

$w(y, v_y; \Psi_\tau)$ is simply related to $w_{\mathscr{X},\mathscr{Y}}(x, y; \Psi_\tau)$, which is calculable in the standard way:

$$w_{\mathscr{X},\mathscr{Y}}(x, y; \Psi_\tau) = \int_{-\infty}^{\infty} |\Psi(x, y, z; \tau)|^2 dz = \left(\frac{m\omega}{4\pi\hbar}\right) \left| \sqrt{\left(\frac{m\omega}{2}\right)} \varphi\left(\frac{m\omega x}{2}\right) \right|^2$$

Since $v_y = -(\omega x/2)$ at τ, the density $w(y, v_y; \Psi_\tau)$ is given by

$$w(y, v_y; \Psi_\tau) = \frac{2}{\omega} w_{\mathscr{X},\mathscr{Y}}\left(-\frac{2v_y}{w}, y; \Psi_\tau\right) = \left(\frac{m\omega}{4\pi\hbar}\right) m|\varphi(-mv_y)|^2$$

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 previously been emphasized by one of us (Margenau, 1963). The above expression for $w(y, v_y; \Psi_\tau)$ shows that such independence of $\mathscr{M}_1(\mathscr{Y})$ and $\mathscr{M}_1(\mathscr{V}_y)$ actually obtains for this particular $\mathscr{M}_2(\mathscr{Y}, \mathscr{V}_y)$. To see this more clearly, note that

$$w(y, v_y; \Psi_\tau) = w_{\mathscr{Y}}(y)\, w_{\mathscr{V}_y}(v_y)$$

where $w_{\mathscr{Y}}$ and $w_{\mathscr{V}_y}$ are computed in the usual quantum mechanical manner:

$$\begin{aligned} w_{\mathscr{Y}}(y) &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dz\, |\Psi(x, y, z; \tau)|^2 \\ &= \left(\int_{-\infty}^{\infty} dx \left| \sqrt{\left(\frac{m\omega}{2}\right)} \varphi\left(\frac{m\omega x}{2}\right) \right|^2 \right) \left(\int_{-\infty}^{\infty} dz\, |\zeta(z)|^2 \right) \left(\frac{m\omega}{4\pi\hbar}\right) \\ &= \frac{m\omega}{4\pi\hbar} \end{aligned}$$

$$\begin{aligned} w_{\mathscr{V}_y}(v_y) &= \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dp_{y_0} \int_{-\infty}^{\infty} dz_0\, |\langle \eta_{x_0 p_{y_0} z_0}, \Psi_\tau \rangle|^2 \delta\left[v_y - \left(\frac{p_{y_0}}{m} - \frac{\omega x_0}{2} \right) \right] \\ &= \delta(0)\, m|\varphi(-mv_y)|^2 \text{ `=' } m|\varphi(-mv_y)|^2 \end{aligned}$$

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 $\mathscr{M}_1(\mathscr{Y})$ and $\mathscr{M}_1(\mathscr{V}_y)$ performed via the present $\mathscr{M}_2$ yield *stochastically independent* results. But of course the uncertainty principle still holds; indeed here $\Delta\mathscr{Y}\,\Delta\mathscr{V}_y \to \infty$, provided the Δ's are standard deviations related to collectives of simultaneous $\mathscr{Y}$, $\mathscr{V}_y$-measurements.

In conclusion, we summarize the salient results obtained:

(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 seems to be the only logical foundation 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 contradictions noted under (1).

(5) If quantum theory entails only a *weak correspondence* between operators and observables, i.e. if it is assumed only that Hermitean 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 substantial 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 a certain simple (type A) theory of measurement to describe simultaneous measurement processes; nevertheless, quantum mechanically certified simultaneous measurement schemes of the historical kind (here called type B) can be developed.

(8) Derivation and comparison of the joint probability distribu-

† 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, 1950, pp. 369–375), such procedures have been called 'successful' measurements.) Occasionally one sees discussion 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 (Arthurs & Kelly, 1965; She & Heffner, 1966; Prugovečki, to be published) 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'.

tions associated with several methods of simultaneous measurement reveal 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 indicates that quantal joint probabilities for noncommuting observables are probably not functionals of state alone but depend as well on 'historical' factors concerning particular $\mathscr{M}_2$'s.

References

Arthurs, E. and Kelly, J. L., Jr. (1965). *Bell System Technical Journal*, **44**, 725.

Birkhoff, G. and von Neumann, J. (1936). *Ann. Math.* **37**, 823.

Bohr, N. (1949). In: *Albert Einstsein: Philosopher–Scientist*, ed. Schilpp, P., pp. 201–241. Tudor.

Cohen, L. and Margenau, H. (1967). In: *Quantum Theory and Reality* (Ch. 4). Springer-Verlag, 1967.

Cohen, L. (1966). *Journal of Mathematics and Physics*, **7**, 781.

Dirac, P. A. M. (1945). *Review of Modern Physics*, **17**, 195.

Einstein, A., Podolsky, B. and Rosen, N. (1935). *Physical Review*, **47**, 777.

Feynman, R. P. (1965). *Quantum Mechanics and Path Integrals*, pp. 96–98. McGraw-Hill.

Gleason, A. M. (1957). *Journal of Mathematics and Mechanics*, **6**, 885.

Groenewold, H. J. (1946). *Physica's Gravenhage*, **12**, 405.

Heisenberg, W. (1930). *The Physical Principles of Quantum Theory*. Chicago.

Jammer, M. (1966). *The Conceptual Development of Quantum Mechanics*, McGraw-Hill.

Mackey, G. W. (1963). *Mathematical Foundations of Quantum Mechanics*, p. 70. Benjamin.

Mandl, F. (1957). *Quantum Mechanics*. Academic Press.

Mandl, F. (1959). *Introduction to Quantum Field Theory*. Interscience.

Margenau, H. (1937). *Philosophy of Science*, **4**, 352–356.

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

Margenau, H. (1954). *Physics To-day*, **7**, 6.

Margenau, H. and Hill, R. N. (1961). *Progress of Theoretical Physics*, **26**, 727.

Margenau, H. (1963a). *Annals of Physics (U.S.A.)*, **23**, 469.

Margenau, H. (1963b). *Philosophy of Science*, **30**, 6.

Moyal, J. E. (1949). *Proceedings of the Cambridge Philosophical Society*, **45**, 99.

von Neumann, J. (1931). *Annals of Mathematics*, **32**.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. English translation by R. T. Beyer, p. 351. Princeton.

Park, J. L. (1968a). *American Journal of Physics*, **36**, 211.

Park, J. L. (1968b). 'Quantum Theoretical Concepts of Measurement.' *Philosophy of Science* (to be published).

Peierls, R. (1935). *Nature, London*, **135**, 957.

Piron, C. (1964). *Helvetica physica acta*, **37**, 439.

Prugovečki, E. 'On a Theory of Measurement of Incompatible Observables in Quantum Mechanics' (to be published).

Reichenbach, H. (1944). *Philosophic Foundations of Quantum Mechanics*, University of California.
Roman, P. (1965). *Advanced Quantum Theory*, Addison-Wesley.
Segal, I. E. (1947). *Annals of Mathematics*, **48**, 930–948.
She, C. Y. and Heffner, H. (1966). *Physical Review*, **152**, 1103.
Shewell, J. R. (1959). *American Journal of Physics*, **27**, 16.
Streater, R. F. and Wightman, A. S. (1964). *PCT, Spin and Statistics and All That*, p. 100. Benjamin.
Suppes, P. (1961). *Philosophy of Science*, **28**, 378; and in *Philosophy of Science*: The Delaware Seminar, v. 2.
Suppes, P. (1966). *Philosophy of Science*, **33**, 14–21.
Temple, G. (1935). *Nature, London*, **135**, 957.
Varadarajan, V. S. (1962). *Communications on Pure and Applied Mathematics*, **15**, 189–217.
Wick, G. C., Wightman, A. S. and Wigner, E. P. (1952). *Physical Review*, **88**, 101.