Progress of Theoretical Physics, Vol. 26, No. 5, November 1961

# Correlation between Measurements in Quantum Theory

Henry MARGENAU* and Robert Nyden HILL

*Yale University, New Haven, Conn., U.S.A.*

(Received July 15, 1961)

An attempt to extend the postulational basis of quantum theory by introducing correlations between the results of measurements, developed in this article, leads to negative joint probabilities for otherwise meaningful sets of measured values. Since, so far as can be seen, the attempt made is the only one compatible with the theory of random variables, and is incompatible with the structure of Hilbert space, we conclude that correlations are absent. This result, though it is tantamount to a denial of von Neumann's projection postulate and the "reduction of wave packets" on measurement, is nevertheless shown to be entirely satisfactory from the physical point of view.

## § 1. Introduction

It has become customary in textbooks and courses on quantum mechanics to allege that a measurement of an observable whose operator is $X$, when performed on a physical system in a state $\psi$, will leave the system in the eigenstate $\psi_{x_i}$ which corresponds to the measured value $x_i$. If $\psi$ is regarded as a vector in Hilbert space, the measurement is represented by an operator $M$ which, when acting on $\psi$, projects it in the direction of $\psi_{x_i}$; the index $i$ is unpredictable on the basis of the usual axioms of the quantum theory, although the change in state can be formalized by assigning to $M$ a definite mathematical operator. This proposition, seemingly first advanced by von Neumann,[1] will be called the *projection postulate*.

The context in which it arose is an interesting one. There are, he suggests, three stages of causal or acausal description. The first renders measurement a totally statistical act, allowing only the assignment of relative frequencies to measured values without implying what value a second measurement, performed immediately after the first, will yield. The second stage permits a variance of measured values only before an initial measurement, which then forces the outcome of later measurements of the same observable in strict deterministic fashion. According to the third or highest version of causality, every measurement is completely determined, as it is in classical physics. The second stage, von Neumann holds, is the one upon which quantum mechanics operates, and to insure it he introduces the projection postulate.

---

* Fulbright Visiting Professor at the University of Tokyo, during the summer semester 1961.

Now it is a curious fact that in the applications of quantum mechanics to any physical situation this postulate is never needed, even though it is occasionally invoked—uselessly to be sure—in the explanation of cloud chamber tracks.[2] Among its less desirable consequences are the Einstein-Podolski-Rosen paradox[3] and the so-called reduction of the wave packet. The latter phenomenon, which is the contraction of a wave packet spread through space into a delta function when a precise position measurement is made, has given rise to the query whether quantum mechanics conveys an objective description of physical reality or is merely a subjective vehicle for an observer's knowledge. Such grave implications, together with important recent reconsiderations[4] of the measurement problem, have motivated the present analysis which tends to show an inconsistency of a formal sort in the unqualified acceptance of the projection postulate. While our discussion will be limited to elementary, non-relativistic quantum theory, no obstacle is apparent which would preclude its extension to four-dimensional analysis and to field theory.

Aside from the fact that the projection postulate is almost never included among the formal axioms of the theory, its use being limited to collateral expositions of a vague philosophic kind and of no apparent relevance to physical problems, it invites fundamental suspicion for the following reason. The so-called "Copenhagen interpretation" is said to be a radically statistical one, permitting no connection with the prediction-potent, individual events of classical physics. Now in every normal theory of probability or random variables it is understood that a single observation cannot establish a probability distribution, a large sample of observations being required for this end. Yet the customary quantum theory, if it includes the postulate in question, claims to be a radical probability theory wherein "God plays dice" and at the same time pretends to perform the most unorthodox feat of creating, in a single act of measurement, knowledge of an entire probability distribution, thus setting its competence high above what is regarded as normal by statisticians. It is for this reason among others that one of us has in the past directed criticism against the projection postulate (understood in von Neumann's sense),[5] and the strictures were based upon the observation that there are many good measurements, fully accredited by physicists, which empirically violate the postulate.[6] In the present article an attempt is made to incorporate the postulate in the foundations of the quantum theory. The success is only partial, but the effort is interesting inasmuch as it shows that negative probabilities are introduced when the requisite axiomatic extension is made.

A set of basic propositions sufficient for all of quantum mechanical analysis is the following.[7]

1. To every observable there corresponds an operator.
2. The operand $\psi$ of the operator represents a physical state.
3. The numerical values which a measurement upon an observable with oper-

ator $P$ can yield are the eigenvalues of $P$.

4. When a physical system is in a state $\psi$, the expectation value of a sequence of measurements on the observable whose operator is $P$ is given by

$$E(P)\equiv\bar{p}=\int\psi^{*}P\psi\,d\tau\equiv(\psi^{*},\,P\psi).$$

This axiom is written in the coordinate (Schrödinger) representation; its isomorphic forms in other representations are well known.[8)]

5. States change in time according to the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t}=H\psi$$

or its equivalents in other representations. Added to this list are symmetry requirements, like the exclusion principle, which are of perhaps a less fundamental or even derivative status. They will not be included here. Perhaps it should be said, however, that a separate premise concerning single-valuedness and normalizability of states is quite unnecessary, since such items follow from 4. The uncertainty principle, of course, is implied by the set when 1 is implemented by a specific assignment of operators which are empirically found to be satisfactory. Finally, 4 can be shown to entail the statement: $|(\psi^{*},\psi_{x_i})|^2$ is the probability that the value $x_i$ (whose correlated eigenstate is $\psi_{x_i}$) will occur in a measurement of $X$ on a system in state $\psi$.

## § 2. Covariance among measurements

The preceding axioms are in complete conformity with the normal procedures in random-variable analysis.[9)] What they say ignores correlations between the results of different measurements. On the other hand, the projection postulate stipulates a very specific correlation between measurements, and it cannot be discussed unless our set is suitably enlarged to provide for correlations. What one would like to know is the probability that a joint measurement of two random variables, $X$ and $Y$, shall yield values $x_i$ and $y_j$, provided the system is known to be in some well-defined quantum state $\psi$. The latter will be taken for simplicity to be a pure case, not a mixture. By joint measurement we mean either two simultaneous measurements of $X$ and $Y$ or, as an interesting subclass of such procedures, the measurement of $X$ at time $t_1$ and of $Y$ at time $t_2$. The variables $X$ and $Y$ can of course be identical.

This specification contradicts the claim, often encountered in the literature, that certain variables, like position and momentum, cannot be measured simultaneously. If this were true, such measurements must be ruled out either by the axioms of quantum mechanics or by the empirical contingencies of experimental physics. The latter is certainly not the case, and the former is impossible because the axioms fail to speak of correlations. The joint probability in question, as defined in the foregoing paragraph, will be designated as

$$P(x_i, y_j ; \psi).$$

Briefly, it is the probability for the occurrence of $x_i$ and $y_j$ on the evidence $\psi$. It must satisfy the relations

$$\sum_{ij} P(x_i, y_j ; \psi)=1, \tag{1}$$

$$\sum_i P(x_i, y_j ; \psi)=P(y_j ; \psi), \ \sum_j P(x_i, y_j ; \psi)=P(x_i ; \psi). \tag{2}$$

The ordinary probabilities, $P(x_i ; \psi)$ and $P(y_j; \psi)$ are indeed given by our axioms, viz:

$$P(x_i ; \psi)=|(\psi^*, \psi_{x_i})|^2 \tag{3}$$

with a similar relation for $P(y_j ; \psi)$.

The quantity thus introduced must be carefully distinguished from the probability $W(x_i, y_j)$ that, when $x_i$ is known to occur with certainty, $y_j$ will ensue. It, too, is given by the axiom set:

$$W(x_i, y_j)=|(\psi^*_{x_i}, \psi_{y_j})|^2. \tag{4}$$

If the projection postulate were valid, the following relation might be thought to hold:

$$P(x_i, y_j ; \psi)=|(\psi^*, \psi_{x_i})|^2 W(x_i, y_j), \tag{5}$$

provided $y_j$ is observed immediately after $x_i$. This version, however, cannot be accepted because it violates one of the relations (2).

A possible definition of $P(x_i, y_j ; \psi)$, indeed one which is compatible with and almost suggested by axioms 1–5, is this:

$$P(x_i, y_j ; \psi)=|(\psi^*, \psi_{x_i})(\psi, \psi^*_{y_j})|^2. \tag{6}$$

The closure relation for complete sets of eigenstates may be used to show that this formulation satisfies relations 1, 2 and 3, but it is clearly different from Eq. (5) and it does not allow measurements on $Y$ to depend in any way on the outcome of measurements on $X$. For this reason, one might look upon Eqs. (5) and (6) as imperfect and search for another possible definition of $P$, perhaps one closer to Eq. (5), at any rate one which introduces correlations. This is best done via the concept of covariance, as follows:

Let $X$ and $Y$ be random variables. Then

$$\mathrm{Cov}(XY)=\mathrm{E}(XY)-\bar{x}\bar{y}. \tag{7}$$

If $X$ and $Y$ are identical, $\mathrm{Cov}(XY)$ becomes the variance of $X$, $\mathrm{Var}(X)$. The structure of our axioms, particularly 4, suggests at once that we postulate for a quantum mechanical ensemble of measurements on a state $\psi$

$$\mathrm{Cov}(XY)=(\psi^*, XY\psi)-(\psi^*, X\psi)(\psi^*, Y\psi). \tag{8'}$$

This interpretation differs from the one which (illegitimately, it seems) takes the operator $XY$ to be associated with an observation of $X$ followed immediately by an observation of $Y$.

Assuming that $X$ and $Y$ are hermitean, one can at once establish the following reasonable facts.

If $\psi$ is an eigenstate of $X$

$$\mathrm{Cov}(XY)=0,$$

for in that instance

$$(\psi^*_{x_i}, XY\psi_{x_i})=(X^*\psi^*_{x_i}, Y\psi_{x_i})=x_i\bar{y}=\bar{x}\bar{y}.$$

The same result holds when $\psi$ is an eigenstate of $Y$.

In general, however, $\mathrm{Cov}(XY)$ according to Eq. (8′) is complex. But

$$\mathrm{Cov}^*(XY)=\mathrm{Cov}(YX).$$

It is therefore indicated that Eq. (8′) be replaced by the definition

$$\mathrm{Cov}(XY)=\left(\psi^*, \frac{XY+YX}{2}\psi\right)-(\psi^*, X\psi)(\psi^*, Y\psi). \tag{8}$$

We shall thus base our further inquiry upon this formulation, which leaves the conclusions concerning eigenstates unaltered.

A further interesting relation can be easily established. If, in one of the well-known forms of Schwarz' inequality, to wit

$$(f^*, f)(g^*, g)\geq\frac{1}{4}[(f^*, g)+(f, g^*)]^2$$

we take

$$f=(X-\bar{x})\psi,$$
$$g=(Y-\bar{y})\psi$$

we find at once

$$[\mathrm{Cov}(XY)]^2\leq\mathrm{Var}(X)\cdot\mathrm{Var}(Y). \tag{9}$$

In consequence one is permitted to define a *correlation coefficient* $\sigma$ after the manner of random-variable theory:

$$\sigma(X, Y)=\mathrm{Cov}(XY)[\mathrm{Var}(X)\,\mathrm{Var}(Y)]^{-1/2} \tag{10}$$

and the range of $\sigma$ is the expected one, $-1\leq\sigma\leq1$.

A few examples illustrating these relations seem to be in place.

a) The spin components of a single electron "anticommute". Hence, if $X$ and $Y$ are $x$- and $y$-components of the spin,

$$\mathrm{Cov}(XY)=-\bar{x}\bar{y}=(a^*b+ab^*)(ia^*b-iab^*)\frac{\hbar^2}{4},$$

provided the spin function is $\begin{pmatrix} a \\ b \end{pmatrix}$. For a state $\psi$ which is an eigenstate of the $z$-component, either $a$ or $b$ is zero, and the covariance vanishes.

b) Less trivial, perhaps, is the case of a system with angular momentum $J=\hbar$, we now write $X$, $Y$, $Z$ for the components of $\boldsymbol{J}/\hbar$. These operators have the matrix form:

$$X=\sqrt{\frac{1}{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad Y=\sqrt{\frac{1}{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad Z=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Their common eigenvalues are $-1$, 0 and 1 and their eigenvectors form the following list, to which reference will be made again in the next section. The

| Eigenvalue | $-1$ | 0 | 1 |
|---|---|---|---|
| $\psi_x=$ | $\begin{pmatrix} \frac{1}{2} \\ -\sqrt{\frac{1}{2}} \\ \frac{1}{2} \end{pmatrix}$ | $\begin{pmatrix} \sqrt{\frac{1}{2}} \\ 0 \\ -\sqrt{\frac{1}{2}} \end{pmatrix}$ | $\begin{pmatrix} \frac{1}{2} \\ \sqrt{\frac{1}{2}} \\ \frac{1}{2} \end{pmatrix}$ |
| $\psi_y=$ | $\begin{pmatrix} \frac{1}{2} \\ -\sqrt{\frac{1}{2}}i \\ -\frac{1}{2} \end{pmatrix}$ | $\begin{pmatrix} \sqrt{\frac{1}{2}} \\ 0 \\ \sqrt{\frac{1}{2}} \end{pmatrix}$ | $\begin{pmatrix} \frac{1}{2} \\ \sqrt{\frac{1}{2}}i \\ -\frac{1}{2} \end{pmatrix}$ |
| $\psi_z=$ | $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ | $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ | $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ |

operator

$$\frac{1}{2}(XY+YX)=\frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

has eigenvalue $-1$, 0 and 1. Simple computation shows that, for a state

$$\psi=\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$\mathrm{Cov}(XY)=-\frac{1}{2}(a^*c-ac^*)i+\frac{1}{2}[(a^*+c^*)b+(a+c)b^*] \times[(a^*-c^*)b-(a-c)b^*]i.$$

Interestingly again, if $\psi$ is an eigenvector of $Z$, the covariance vanishes. The same is true, of course, when $\psi$ is one of the $\psi_x$ or $\psi_y$ in accordance with theorem a) following Eq. (8′).

c) Next, let us compute the covariance between position measurements at time $t=0$ and at time $t$. The operator $X$ corresponding to a measurement of $x$ at $t=0$ is simply $x\cdot$; at the later time it is

$$X_t \sim T^* xT,$$

where $T$ is the time-development operator, $T=\exp\left(-i\int\frac{H}{\hbar}dt\right)$. For $H$ we shall take the Hamiltonian of a free particle. In that case, since for any operators $A$ and $B$

$$e^A Be^{-A}=B+[A, B]+\frac{1}{2!}[A, [A, B]]+\cdots$$

in the notation of commutator brackets, we have

$$X_t=e^{\frac{i}{\hbar}Ht}xe^{-\frac{i}{\hbar}Ht}=x+\frac{it}{\hbar}[H, X]=x+\frac{t}{m}P \tag{12}$$

where $P$ is the linear momentum operator. In general $T$ satisfies

$$T^{-1}=T^*$$

when it is written in the form of a differential operator; in matrix form it is unitary. If the eigenfunctions of an operator $Q$ at $t=0$ are $\psi_i$, then the eigenfunctions of the same operator at time $t$ are $T^*\psi_i$. To see this, assume the equation

$$Q\psi_i=q_i\psi_i$$

to be satisfied at $t=0$. Let $\psi_i^t$ be the solution of

$$(T^*QT)\psi_i^t=q_i^t\psi_i^t.$$

On multiplying by $T$ from the left we obtain

$$Q(T\psi_i^t)=q_i^t(T\psi_i^t)$$

which means that $T\psi_i^t$ is an eigenfunction and $q_i^t$ an eigenvalue of $Q$. Hence $q_i^t$ must equal $q_i$; the eigenvalues cannot change in time. And since $T\psi_i^t=\psi_i$, we have $\psi^t=T^*\psi_i$ in view of Eq. (11). Evidently, $\psi_i^t$ is the ancestor of $\psi_i$ at time $-t$ in accordance with the Schrödinger equation.

On using Eqs. (8) and (12) we obtain, upon expansion,

$$\mathrm{Cov}(XX_t)=\mathrm{Var}(X)+\frac{t}{m}\mathrm{Cov}(XP) \tag{13}$$

where $X$ without subscript refers to $t=0$.

For $\psi$ we assume a packet of the form

$$\psi=(\pi a^2)^{-1/4}\exp\left(ikx-\frac{x^2}{2a^2}\right) \tag{14}$$

at $t=0$, which transforms itself into

$$\psi_t=(\pi a^2)^{-1/4}\left(1+\frac{i\hbar t}{ma^2}\right)^{-1/2}\exp-\left[\frac{x^2-2ia^2kx+\dfrac{ia^2k^2\hbar t}{m}}{2\left(a^2+\dfrac{i\hbar}{m}t\right)}\right] \tag{15}$$

at time $t$. It is then easily seen that

$$\mathrm{Cov}(XP)=0 \tag{16}$$

so that

$$\mathrm{Cov}(XX_t)=\mathrm{Var}(X)=\frac{a^2}{2} \tag{17}$$

at all times.

The correlation coefficient defined in Eq. (10), however, depends on $t$. $\mathrm{Var}(X)$ has just been computed, and

$$\mathrm{Var}(X_t)=\mathrm{Var}\left(X+\frac{t}{m}P\right)=\mathrm{Var}(X)+\frac{t^2}{m^2}\mathrm{Var}(P)+\frac{2t}{m}\mathrm{Cov}(XP).$$

But $\mathrm{Var}(P)=\hbar^2/2a^2$, hence, in view of Eqs. (16) and (17),

$$\mathrm{Var}(X_t)=\frac{a^2}{2}\left[1+\left(\frac{\hbar t}{ma^2}\right)^2\right].$$

Therefore the correlation coefficient between a position measurement at time 0 and time $t$ is

$$\sigma=\left[1+\left(\frac{\hbar t}{ma^2}\right)^2\right]^{-1/2}.$$

It remains near 1 if the initial extent of the wave packet, $a$, is large. The fact that it does not depend on $k$, the speed of the packet, is also reasonable, for $\sigma$ is 1 for random variables which are connected by a linear relation, e.g. $x_t=x_0+vt$.

## § 3. Joint probabilities

Having shown that the covariance defined in Eq. (8) is not without attractive features, we next examine the joint probabilities it entails. As is known from statistics[9)]

$$\mathrm{Cov}(XY)=\sum_{ij}P(x_i y_j)x_i y_j-\bar{x}\bar{y}. \tag{18}$$

Eq. (8) takes this form if we use the alternate expansions

$$\psi=\sum_i a_i u_i \quad \text{and} \quad \psi=\sum_j b_j v_j$$

in orthonormal sets $u$ and $v$ which are, respectively, eigenstates of $X$ and $Y$. For then

$$(\psi^*, XY\psi)=\int X^*\sum_\lambda a_\lambda^* u_\lambda^* Y\sum_\mu b_\mu v_\mu d\tau=\sum_{ij}\left(a_i^* b_j\int u_i^* v_j d\tau\right)x_i y_j.$$

Comparison with Eqs. (18) and (8) allows a simple identification of $P$:

$$P(x_i, y_j; \psi)=R[a_i^* b_j \int u_i^* v_j d\tau]=R[(\psi^*, \psi_{x_i})(\psi, \psi_{y_j}^*)(\psi_{x_i}^*, \psi_{y_j})]. \qquad (19)$$

Because of the closure relation these joint probabilities satisfy Eqs. (1), (2) and (3). If $\psi$ is either $\psi_{x_i}$ or $\psi_{y_j}$, they also lead to Eq. (4), but they are not in general accord with the projection postulate in the narrow form here under consideration. Yet, as is clear from the preceding section, they introduce correlations between measurements. When $X$ and $Y$ commute

$$P(x_i, y_j; \psi)=P(x_i, \psi)\cdot\delta_{ij},$$

provided proper allowance is made for degeneracies. Most unfortunately, however, the defining relation (19) does not prevent the joint probabilities from being negative. This is a rather essential aspect connected with the present attempt to introduce correlations between measurements, and it is difficult to see how it can be avoided; its causes lie very deep, and their removal would require some radical changes in the axioms listed in the Introduction. For such reasons it seemed worthwhile to illustrate the case by computing a few examples.

a) Returning to the system with angular momentum $J=1$ to which discussion b) of § 2 was devoted, we label the eigenvalues $-1$, 0 and 1 by indices $\rho$ and $\sigma$, allowing them the values $-1$, 0 and 1. The state $\psi$ has again the form of a column vector with components $a$, $b$ and $c$. For convenience we introduce the abbreviations

$$s=\sqrt{\frac{1}{2}}(a+c), \quad d=\sqrt{\frac{1}{2}}(a-c).$$

Now

$$P(x_i, y_j; \psi)\equiv P_{\rho\sigma}=R(\psi^+, u_\rho)(v_\sigma^+, \psi)(u_\rho^+, v_\sigma),$$

$u_\rho$ and $v_\sigma$ being the eigenstates $\psi_x$ and $\psi_y$ listed in 2b. One finds

$$(\psi^+, u_{-1})=\sqrt{\frac{1}{2}}(s^*-b^*),$$

$$(\psi^+, u_0)=d^*,$$

$$(\psi^+, u_1)=\sqrt{\frac{1}{2}}(s^*+b^*),$$

$$(v_{-1}^+, \psi)=\sqrt{\frac{1}{2}}(d+ib),$$

$$(v_0^+, \psi)=s,$$

$$(v_1^+, \psi)=\sqrt{\frac{1}{2}}(d-ib).$$

The last factor of $P_{\rho\sigma}$ is

$$(u_\rho^+, v_\sigma) = \begin{array}{c|ccc} \rho \backslash \sigma & -1 & 0 & 1 \\ \hline -1 & +\frac{1}{2}i & \sqrt{\frac{1}{2}} & -\frac{1}{2}i \\ 0 & \sqrt{\frac{1}{2}} & 0 & \sqrt{\frac{1}{2}} \\ 1 & -\frac{1}{2}i & \sqrt{\frac{1}{2}} & +\frac{1}{2}i. \end{array}$$

From these all joint probabilities can be compounded. In the case where $\psi$ is an eigenstate of $J_z$ (see § 2b for the form of these eigenstates!) one obtains

$$P_{\rho\sigma} = \begin{vmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{vmatrix} \quad \text{if } J_z = -1 \text{ or } +1$$

and

$$P_{\rho\sigma} = \begin{vmatrix} \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \end{vmatrix} \quad \text{if } J_z = 0.$$

These are perfectly acceptable probability distributions whose meaning is clear and simple. For instance, if $J_z$ is known *a priori* to be 1 or $-1$, there is no chance of finding $J_x$ and $J_y$ manifesting equal values in simultaneous measurements, nor can they exhibit opposite values. One can only obtain a zero value for one and $+1$ or $-1$ for the other. If $J_z$ is 0, the situation is the reverse. Notably, negative probabilities do not occur.

Generally, however, the situation is quite different. If $s$, $d$ and $b$ (which must obey the normalization rule $|s|^2+|d|^2+|b|^2=1$) are taken to be, respectively, 0.6, 0.6 and 0.53, one calculates

$$P_{-1-1} = \frac{1}{4} R\{i(s^*-b^*)(d+ib)\} = -0.009.$$

b) Let $x_i$ be the position $x$ of a particle, $y_j$ its simultaneously measured linear momentum, so that (*if* $p = k\hbar$).

$$P(x, k; \psi) = R\int \delta(q_1 z)\,\psi^*(q_1)\,dq_1 \int c e^{-ikq_2}\psi(q_2)\,dq_2 \int \delta(q_3, x)\,c^* e^{ikq_3}\,dq_3$$

$$= |c|^2 R\{\psi^* e^{ikx} \int \psi(q)\,e^{-ikq}\,dq\}.$$

Assume for $\psi$ a wave packet of the form (14). Evaluation then leads to the expression

$$P(x, k; \psi) = \sqrt{2}\,|c|^2 \cos(k-k_0)\,x \exp -\frac{1}{2}\left[\frac{x^2}{a^2} + (k-k_0)^2 a^2\right]. \quad (20)$$

The fact that $P$ has an absolute maximum at $x=0$ and $k=k_0$ is satisfying, of course, as is the diffusive exponential decrease with $x$ and $k$; but the oscillating factor prevents acceptance of the result.

c) As a last example, we discuss the joint probability, again in the simple case of a moving free particle, for position measurements at different times. Here we have

$$P(x_1, x_2{}^t; \psi) = R\{(\psi^*, \psi_{x_1})(\psi, \psi_{x_2}^{t*})(\psi_{x_1}^* \psi_{x_2}^t)\}, \quad (21)$$

$$\psi_{x_1} = \delta(x, x_1),$$

$$\psi_{x_2}^t = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[ik(x-x_2) + i\frac{\hbar t}{2m}k^2\right] = \frac{1-i}{2}\left(\frac{m}{\pi\hbar t}\right)^{1/2} \exp\left[-i\frac{m}{2\hbar t}(x-x_2)^2\right].$$

Insertion in Eq. (21) gives (with certain precautions when $t\to 0$)

$$P(x_1 x_2{}^t, \psi) = R\left\{\frac{m}{ht}\psi^*(x_1)\exp\left[-i\frac{m}{2\hbar t}(x_1-x_2)^2\right]\right.$$

$$\left.\times \int \psi(x)\exp\left[i\frac{m}{2\hbar t}(x-x_2)^2\right]dx\right\}.$$

Here we use once more the form (14) for $\psi$. We carry out the integration and, to make the result less unwieldy, put $x_1=0$, obtaining

$$P(0, x_t; \psi) = R\left\{\frac{\sqrt{2}\,m}{ht}\left(1-\frac{ima^2}{\hbar t}\right)^{-1/2} \exp -\frac{1}{2}\left[\frac{a^2\left(k-\frac{mx}{\hbar t}\right)^2}{1-\frac{ima^2}{\hbar t}}\right]\right\}. \quad (22)$$

Our choice of normalization has been such that $p$ now represents a probability density (per unit range of $x_1$ and $x_2$).

Two limiting cases are of primary interest. For large $t$, i.e. if $\hbar t/m \gg a^2$,

$$P(0, x^t; \psi) \approx \sqrt{2}\,\frac{m}{ht}\exp\left[-\frac{a^2}{2}\left(k_0 - \frac{mx}{\hbar t}\right)^2\right].$$

This is always positive and has a maximum at the classical position of the particle, $x = k_0\hbar/m\cdot t$.

In the opposite extreme, when $\hbar t/m \ll a^2$

$$P(0, x^t; \psi) \approx R\left\{\left(\frac{m}{\pi h a^2 t}\right)^{1/2} \exp\left[i\left(\frac{\pi}{4} - \alpha\right)\right]\right\} \tag{23}$$

where

$$\alpha = \left(k - \frac{mx}{\hbar t}\right)^2 \frac{\hbar t}{2m}$$

or

$$P(0, x^t; \psi) \propto \cos\alpha + \sin\alpha.$$

For small $t$, $x$ can be chosen so that this quantity is negative. Closer inspection of Eq. (23) reveals, for finite $k_0$ and $x$, the satisfactory limit

$$\lim_{t\to 0} P(0, x; \psi) \propto \delta\left(x, \frac{k_0 h}{m} t\right).$$

In all preceding examples Eq. (19) was employed as the definition of joint probability. Although reasonable, this definition is not unique. Its lack of uniqueness arises from the circumstance that we have inferred its form from the single product moment, $(\psi^*, XX\psi)$, leaving all higher moments $(\psi^*, X^n Y^m \psi)$ with their possible variants $(\psi^*, X^{n-1} Y^m X\psi)$, etc., unspecified. We have therefore considered another possibility which is also compatible with Eq. (18) and has certain other advantages, namely

$$P(x, y; \psi) = (2\pi)^{-2} \int\int_{-\infty}^{\infty} \rho(\alpha, \beta) \exp[-i(\alpha x + \beta y)]\, d\alpha\, d\beta$$

with

$$\rho(\alpha, \beta) = (\psi^*, \exp[i(\alpha X + \beta Y)]\psi)$$

where in accordance with our previous notation $X$ and $Y$ are operators.[1c] When the analysis is carried through for a case similar to example $b$ in which $X$ is the position and $Y$ the momentum of a particle, but $\psi$ is chosen to be a superpositian of two gaussian functions

$$\psi(x) = [2a\sqrt{\pi}\,(1 + e^{-b^2/a^2})]^{-1/2} [e^{-(x-b)^2/2a^2} + e^{-(x+b)^2/2a^2}],$$

one again obtains a result with a trigonometric factor which can cause $P$ to be negative.

There is, however, a convincing argument which establishes the impossibility of introducing any sensible joint probability distribution that exhibits correlations.

The analysis of the possibility of introducing joint probabilities in quantum mechanics proceeds from the observation that a classical joint probability distribution is completely determined only if a certain set of its moments is specified. In the case of random variables which range over an infinite set of values, specification of all the moments is necessary.

Let us now return to the previous example of angular momentum $\hbar$ and attempt to construct a joint probability distribution $P(x_i, y_i; \psi)$, which we now write for brevity in the form $f(x, y)$, where $X$ and $Y$ are taken to be $1/\hbar J_x$ and $1/\hbar J_y$. The eigenvalues of $X$ and $Y$ are then 1, 0, and $-1$. It follows from postulate 3 that we need to consider only ($f(1, \pm 1), f(-1, \pm 1), f(0, \pm 1)$, $f(\pm 1, 0)$ and $f(0, 0)$—i.e. the space of our random variables has only 9 points, so that only nine moments are needed to detemine $f(x, y)$. These moments are $\langle x\rangle$, $\langle y\rangle$, $\langle x^2\rangle$, $\langle y^2\rangle$, $\langle xy\rangle$, $\langle x^2y\rangle$, $\langle xy^2\rangle$, $\langle x^2y^2\rangle$, and the normalization $\langle 1\rangle$. The classical formulae

$$\langle x^n y^m\rangle = \sum_{xy} f(x, y)\, x^n y^m \quad m=0, 1, 2 \text{ and } n=0, 1, 2$$

for these nine moments can be inverted to give

$$\begin{aligned}
f(1, \pm 1) &= 1/4\,(\pm\langle xy\rangle \pm \langle x^2y\rangle + \langle xy^2\rangle + \langle x^2y^2\rangle),\\
f(-1, \pm 1) &= 1/4\,(\mp\langle xy\rangle \pm \langle x^2y\rangle - \langle xy^2\rangle + \langle x^2y^2\rangle),\\
f(0, \pm 1) &= 1/2\,(\pm\langle y\rangle + \langle y^2\rangle \mp \langle x^2y\rangle - \langle x^2y^2\rangle),\\
f(\pm 1, 0) &= 1/2\,(\pm\langle x\rangle + \langle x^2\rangle \mp \langle xy^2\rangle - \langle x^2y^2\rangle),\\
f(0, 0) &= 1 + \langle x^2y^2\rangle - \langle x^2\rangle - \langle y^2\rangle.
\end{aligned}$$

Our previous treatment left all moments with $n$ or $m > 1$ undefined.

We now examine all possible quantum mechanical definitions of the moments. Moments of the form $\langle x^n\rangle$ or $\langle y^n\rangle$ are already specified as $(\psi^*, X^n\psi)$ or $\langle\psi^*, Y^n\psi\rangle$ by the axioms of quantum mechanics. The other moments ("cross moments") will be taken, in similar fashion, to be the expectation values of hermitean operators. These operators will be homogeneous polynomials in $X$ and $Y$, and must reduce to $X^n Y^m$ for commuting operators. Furthermore, symmetry between $X$ and $Y$ must be maintained—i.e. $\langle x^n y^m\rangle$ must go over into $\langle x^m y^n\rangle$ on interchange of $X$ and $Y$. The following choices are then possible with $\alpha$, $\beta$, $\gamma$, $\delta$, and $\varepsilon$ as arbitrary real parameters.

$$\begin{aligned}
\langle xy\rangle &= 1/2(\psi^*, XY + YX\psi),\\
\langle x^2y\rangle &= 1/(2\alpha+\beta)\,(\psi^*, |\alpha(X^2Y + YX^2) + \beta XYX|\psi),\\
\langle xy^2\rangle &= 1/(2\alpha+\beta)\,(\psi^*, |\alpha(XY^2 + Y^2X) + \beta YXY|\psi),\\
\langle x^2y^2\rangle &= 1/(2\gamma+2\delta+2\varepsilon)\,(\psi^*, |\gamma(X^2Y^2 + Y^2X^2) + \delta(XYXY + YXYX)\\
&\qquad + \varepsilon(XY^2X + YX^2Y)|\psi).
\end{aligned}$$

The question of interest is then the following: Do there exist values of these arbitrary parameters for which all of the $f(x, y)$ $(x=-1, 0, 1$ and $y=-1, 0, 1)$ are positive for an arbitrary state function

$$\psi = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \qquad ?$$

Let us assume that such values exist—in particular, that a satisfactory value of $\varepsilon$ exists. In order to find what it should be, we calculate certain of the moments for the arbitrary state:

$$\langle 1\rangle = |a|^2+|b|^2+|c|^2=1, \text{ if } \psi \text{ is normalized}$$
$$\langle x^2\rangle = 1/2(|b|^2+a^*c+ac^*+1),$$
$$\langle xy\rangle = -i/2(a^*c-ac^*),$$
$$\langle xy^2\rangle = \alpha/(2\alpha+\beta)\sqrt{2}[(a^*+c^*)b+(a+c)b^*],$$
$$\langle x^2y\rangle = -i\alpha/(2\alpha+\beta)\sqrt{2}[(a^*-c^*)b-(a-c)b^*],$$
$$\langle x^2y^2\rangle = 1/2(\gamma+\delta+\varepsilon)[2\gamma|b|^2+\varepsilon(|a|^2+|c|^2)].$$

Now the quantities $f(1, 1)+f(1, -1)=1/2(\langle xy^2\rangle+\langle x^2y^2\rangle)$ and $f(1, 0)+f(-1, 0)=\langle x^2\rangle-\langle x^2y^2\rangle$ must be positive or zero for any $\psi$. In particular, for the state $a=-1/\sqrt{2}$, $b=0$, $c=1/\sqrt{2}$, $\langle x^2\rangle=\langle xy^2\rangle=0$ and $\langle x^2y^2\rangle$ must also be zero in order to make the two quantities positive or zero. But for this state, $\langle x^2y^2\rangle=\varepsilon/2(\gamma+\delta+\varepsilon)$, from which we conclude that $\varepsilon$ must be zero. Having fixed the value of $\varepsilon$, we now compute $f(1, 1)$ for the state $a=\exp(i\pi/4)/\sqrt{2}$, $b=0$, $c=\exp(-i\pi/4)/\sqrt{2}$, for which $\langle xy^2\rangle=\langle x^2y\rangle=0$, $\langle x^2y^2\rangle=0$ (as a consequence of $\varepsilon=0$) and $\langle xy\rangle=-1/2$ to give $f(1, 1)=-1/8$. Thus it is not possible to find a satisfactory value of $\varepsilon$, from which we conclude that one cannot define a non-negative $f(x, y)$ which exhibits correlations.

The conclusion that joint probabilities of the usual type cannot be formulated in quantum mechanics was reached by De Broglie,[11] whose reasoning was based, however, entirely on Eq. (5). He regards this equation as fundamental and inevitable.

There arises at this point a rather subtle distinction which is worth noting. The inevitable equation, which relates joint probabilities to so-called conditional probabilities, is this:

$$P(x_i, y_j; \psi) = P(x_i)P'(x_i, y_j) \qquad (5')$$

where the conditional probability $P'$ is the probability that $y_j$ will be observed when $x_i$ is known to be present or, on our specific condition, $P'$ is the relative frequency of $y_j$ in a subclass of measurements all of which yield $x_i$. It is *not* necessarily the same as $W(x_iy_j)$, which represents the probability for the occurrence of $y_j$ when $x_i$ is *known to occur with certainty*, that is, when $\psi$ is an eigenstate of $x_i$. The projection postulate is the statement which says that $P'=W$. De Broglie's argument therefore embraces this postulate; instead of rejecting it, he perhaps redefines joint probabilities in an unusual way.

If, as we suggest, there are no correlations between measurements, $P'(x_i y_j)$ is simply $P(y_j)$ in accordance with Eq. (6), and Eq. (5′) above is satisfied without artificiality, although it is now trivial.

## § 4. Conclusions

There appears to be no reasonable criterion for ruling out those instances in which joint probabilities are negative. Hence it is necessary to reject *in toto* both our definition of the covariance, Eq. (8), and its consequence, such as Eq. (19). The axioms stated in the Introduction do not seem to tolerate this extension. We are thus led back with evident cogency to the proposition which the projection postulate endeavored to avoid, namely to Eq. (6) with its implication that measurements of non-commuting observables are in general uncorrelated. For if (6) is used in Eq. (18), $\mathrm{Cov}(XY)=0$.

By this return, however, nothing of value in quantum mechanics is surrendered. For a certain kind of correlation is automatically carried in the state function itself and becomes manifest in the temporal changes of $\psi$. The structure of the axioms is such that the question as to the connection between a measurement at time $t_1$ and at another time $t_2$ need not be asked. Each independent factor of the definition (6) alone provides the answer to the proper question: What is the probability that $X$ shall have the value $x_i$ at time $t$ if the state is $\psi(t)$? It is, of course, $|(\psi^*(t), \psi_{x_i})|^2$. Employing once more the example of a wave packet, whose form is Eq. (14) at $t=0$, while $\psi(x, t)$ is given by Eq. (15), we find for the probability of encountering the particle at $x$, $t$

$$|(x, t)|^2=\left\{(\pi a^2)\left[1+\left(\frac{\hbar t}{ma^2}\right)^2\right]\right\}^{-1/2}\exp-\left[\frac{\left(x-\frac{k\hbar t}{m}\right)^2}{a^2+\frac{\hbar^2 t^2}{m^2 a^2}}\right]. \qquad (24)$$

This result has all the qualities needed for the description of the motion of particles: It reduces to classical propagation with speed $k\hbar/m$ if either $m$ or $a$ is large; it exhibits quantum mechanical diffusion through the dependence of its half-width on $t$. And it clearly makes no concession to the occurrence of a measurement. If a measurement is made with minimal disturbance, the presumption is that it can be repeated at a later time, yielding a result in accordance with Eq. (24); if it is destructive, repreparation of the state $\psi$ at time 0 must be understood. For example, nothing more than Eq. (24) is needed to explain the production of ion tracks.

There is in fact a bonus in this more modest interpretation. For it preserves reference to the prepared state throughout the process of measurement. If a measurement did produce an eigenstate, the memory of the state before measurement would be lost. An exact measurement of $x$ made upon a state represented by a plane wave would yield a $\delta$-function, and so would a measurement upon a spherical wave. Yet we know that the probability for a second measurement at a later time and a different place will depend on the type of wave that existed before the first measurement was made. This fact is correctly rendered by Eq. (6) and its corollary, Eq. (24), but is denied by the projection postulate.

Earlier, attention was called to von Neumann's three stages of causality. Our conclusion indicates that his threefold division does violence to quantum mechanics and that, insofar as it is tenable at all, we are closer to the first than to the second stage. The logical inconsistency of using a probability theory in which a single observation determines a distribution function is thus removed.

The version of the projection postulate we were forced to deny is the one most widely featured, but probably not the one most generally believed. Thoughtful physicists, when reflecting upon the measurement act, find themselves committed to a statement which is not postulational and is far milder than von Neumann's axiom; it is this. Many types of measurement (not all) involve a *selection* of systems from an original ensemble, and by proper subsequent manipulation one such separated set can then be isolated. This set may represent an eigenstate of the measured observable. One might even wish to argue—against all reasonable practice—that this specification *defines* a quantum mechanical measurement. Even if this argument is accepted there is still no reduction of the wave packet in the sense that the original wave function is suddenly and acausally changed to another form; for the new state, if any, refers to a different physical ensemble for which the old state had no meaning. A simple analogue is the transition from a game involving two dice to one with a single die, for which the spread of probabilities is smaller. This change is not properly discussed as a sudden collapse of probabilities, nor is it puzzling.

A fuller theory of measurement in which all these questions are discussed will be published elsewhere. It describes the measuring act as a conversion from a pure case to a mixture, in von Neumann's sense. This accounts for the so-called "cancelling of phases" said to occur during measurement. The classical probabilities present in the mixture can then, if it is desired, be altered to form new states. Except for the concept of a mixture of states, this complete theory requires no postulate additional to those listed in § 1.

The support of the Guggenheim and National Science Foundations for this work is gratefully acknowledged. We also express our thanks to Professor T. Yamanouchi for helpful comments on this and other problems.

**References**

1) J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, 1932), pp. 111 et seq.
2) See, for example, L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., 1955).

3) H. Margenau, Phys. Rev. **49** (1936), 240.
4) J. L. McKnight, Phil. of Sci. **24** (1957), 321; **25** (1958), 209;
L. Durand, Phil. of Sci. **27** (1960), 115.
5) An acceptable modification, which in essence denies the postulate, is formulated at the end of this article.
6) H. Margenau, Phil. of Sci. **25** (1958), 23; reprinted in C. W. Churchman and P. Ratoosh, *Measurement: Definitions and Theories* (John Wiley and Sons, 1959).
7) This set has not been shown to be consistent or complete in a strict logical sense, a goal which at present eludes our competence. Nor has an attempt ever been made to define the primitive terms or the rules of operation to be employed in the use of the system. Loosely, the latter are the rules of mathematics. The axiomatic system is intuitively felt to be complete, and this feeling is enhanced by the realization that no inconsistency has yet been encountered among its consequences.
8) The abbreviations, $\bar{p}$ for "expectation of P", and the scalar product form $(\phi^*, P, \phi)$ will be employed throughout this paper.
9) The terminology here adopted follows W. Feller, *An Introduction to Probability Theory and its Applications* (John Wiley and Sons, 1957).
10) If $X$ and $Y$ denote positron and momentum operators, the present $P(x, y; \phi)$ becomes Wigner's distribution function. See E. Wigner, Phys. Rev. **40** (1932), 749.
11) L. de Broglie, *Physics and Microphysics* (Harpers, 1960), Appendix.