The quantum probability calculus

Jauch, Joseph-Maria

How to cite

JAUCH, Joseph-Maria. The quantum probability calculus. In: Synthese, 1974, vol. 29, n° 1-4, p. 131–154. doi: 10.1007/BF00484955

This publication URL: https://archive-ouverte.unige.ch/unige:162158

Publication DOI: 10.1007/BF00484955

J. M. JAUCH

THE QUANTUM PROBABILITY CALCULUS*

At bottom, the theory of probability is only common sense reduced to calculation.

Pierre Simon Laplace, 1812

Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means.

Bertrand Russell, 1929

I. INTRODUCTION

Quantum mechanics has opened a vast sector of physics to probability calculus. In fact most of the physical interpretation of the formalism of quantum mechanics is expressed in terms of probability statements.[1]

There are of course large segments of classical physics, too, which are expressed in probabilistic terms. But there is an essential difference between the probabilistic statements of quantum physics and those of classical physics. The present article is devoted to the elucidation of this difference.

The probabilities which occur in classical physics are interpreted as being due to an incomplete specification of the systems under consideration, caused by the limitations of our knowledge of the detailed structure and development of these systems. Thus these probabilities should be interpreted as being of a *subjective* nature.

In quantum mechanics this interpretation of the probability statements has failed to yield any useful insight, because it has not been possible to define an infrastructure whose knowledge would yield an explanation for the occurrence of probabilities on the observational level. Although such theories with 'hidden variables' have been envisaged by many physicists,[2] no useful result has come from such attempts. I therefore take here the opposite point of view which holds that the

Synthese **29** (1974) 131–154.

probabilities in quantum mechanics are of a fundamental nature deeply rooted in the objective structure of the real world. We may therefore call them *objective* probabilities.

It has been noted quite early that the probabilities in quantum theory have some peculiar properties, unrelated to anything previously encountered in classical probability theory. One way of exhibiting these anomalies is by studying joint probabilities for certain pairs of random variables, for instance, those corresponding to the quantum-mechanical position q and the canonically conjugate momentum p.[3] For this case it has been noted by Wigner (1932) already that no positive joint distribution exists.

Various interpretations have been given of this anomaly. I shall not review them critically here, but rather offer yet another one, which I believe corresponds better to the objective character of the quantum probability calculus than previous interpretations.

One point of departure is the observation that the Wigner anomaly for the joint distribution of noncompatible observables is an indication that the classical probability calculus is not applicable for quantal probabilities. It should therefore be replaced by another, more general calculus, which is specifically adapted to quantal systems. In this article I exhibit this calculus and give its mathematical axioms and the definitions of the basic concepts such as probability field, random variable, and expectation values.

Generalized probability calculi have been proposed before.[4] My proposal differs in several respects from previous work on this subject insofar as it is specifically motivated by and adapted to the axiomatic structure of quantum theory as it has been developed by the Geneva School[5] since 1960.

II. PROBABILITY CALCULUS AND PROBABILITY THEORY

The proposed modification of the probability calculus appears more natural if we distinguish between *probability calculus* and *probability theory*.[6] With *calculus* we denote the mathematical formalism devoid of any interpretation of this formalism. With *theory* we refer to the the application of this calculus to various situations involving the occurrences of observable phenomena.

The calculus is a branch of mathematics (in fact of measure theory) and presents no problems of interpretation. The theory on the other hand is beset with numerous difficulties which have been the object of much controversy.

It is remarkable that in none of these controversies was the calculus as such ever questioned and its definitive form as given by Kolmogorov[7] in 1933 has been the basis of all the work on mathematical statistics. The slight generalization of this calculus by Renyi[8] is not essentially different insofar as it removes the restriction of a normalized total probability and replaces it by the basic notion of *conditional* probability.

Little thought has been given to the question why this particular calculus should be so effective in predicting the probabilities of actually occurring events.

The logical situation that we are facing here may be illustrated by an analogy from another branch of mathematics. The discovery of geometry by the Greeks, and in particular its axiomatization by Euclid, led to the idea that the geometry of physical space was unique and absolute. The discovery of non-Euclidean geometries was at first thought to be of no relevance to the geometry of physical space. Only in the physics of the twentieth century, especially through the work of Hilbert and Einstein, did the idea break through that physical geometry is not Euclidean and can actually be determined objectively through physical observations.

Clearly geometry plays the role of the calculus and its interpretation in terms of physical phenomena. It is conceivable that the general theory of relativity could be expressed on the background of a Euclidean space, but in the light of present knowledge it would not be *natural* to do so.

In an analogous way, we contend, it would be possible to express quantum theory on the background of a classical probability calculus, but again, Wigner's work has clearly shown that it would not be natural either to do so.

So just as the geometry of space-time is determined by physical phenomena in the context of a natural theory, it is my belief that probability calculus is equally determined by certain phenomena in the context of quantum theory.

In order to place the new calculus in the proper perspective, I begin with a commentated review of the classical probability calculus.

III. THE CLASSICAL PROBABILITY CALCULUS

The classical calculus of probability is based on a few concepts which I shall introduce and comment briefly in this part. The concepts are: the measurable space, the probability measure, the random variables, the probability distribution function, and the expectation values.

1. *The Measurable Space*

The primary concept of probability calculus is 'the universe of basic events' which in the classical case are identified with a certain class $\mathscr{S}$ of subsets of a set Ω.

The set Ω may be completely arbitrary. Actually as we shall see this set plays in fact only a subsidiary role. What is important are the subsets of the class $\mathscr{S}$ which shall be called the measurable sets.

The class of subsets $\mathscr{S}$ is assumed to be a 'field'. This means it is closed with respect to the operations of the complement, countable unions, and intersections. Furthermore it contains ϕ, the null set, and consequently also Ω, the entire set.

Thus if $S \subset \mathscr{S}$ then the complementary set $S' \in \mathscr{S}$. If S_n $(n=1, 2, \ldots)$ is a countable family of sets from $\mathscr{S}$ then

$$\bigcup_n S_n \in \mathscr{S} \quad \text{and} \quad \bigcap_n S_n \in \mathscr{S}.$$

2. *The Probability Measure*

On the field $\mathscr{S}$ is defined a positive-valued function

$$\mu : \mathscr{S} \to \mathbb{R}^+$$

with the properties

(i) $\mu(\phi)=0; \quad \mu(\Omega)=1.$

(ii) For any pairwise disjoint sequence S_n $(n=1, 2, \ldots)$ such that $S_i' \subset S_k$ for $i \neq k$

$$\mu(\bigcup_n S_n) = \sum_n \mu(S_n) \quad (\sigma\text{-additivity}).$$

This function is the probability measure on $\mathscr{S}$.

We shall refer to the triplet $(\Omega, \mathscr{S}, \mu)$ as the *probability space*. The interpretation of this calculus is that the sets $S \in \mathscr{S}$ denote the possible 'events' and the numbers $\mu(S)$ represent the 'probability' for the occurrence of these events.

3. *Random Variables*

Let $X:\Omega\to\mathbb{R}$ be a real-valued function $X(\omega)$, $\omega\in\Omega$. For any subset $\Delta\in\mathbb{R}$ we denote by

$$X^{-1}(\Delta)=\{\omega \mid X(\omega)\in\Delta\}$$

the *inverse image* of the set Δ under the function X.

A function f is said to be *measurable-B* or simply measurable if for every Borel set $\Delta\in\mathfrak{B}(\mathbb{R})$ the inverse image $X^{-1}(\Delta)\in\mathcal{S}$.

A real *random variable* is a real-valued measurable function on Ω.

It will be seen in the following that the essential property of a random variable, in fact the only property which is really used, is the correspondence which it establishes between Borel sets $\Delta\in\mathfrak{B}(\mathbb{R})$ and the measurable sets. In view of the proposed generalization it is useful to introduce a special notation for this correspondence. Thus we shall denote by $\xi:\mathfrak{B}(\mathbb{R})\to\mathcal{S}$ the correspondence set up by the random variable $X(\omega)$ through

$$X^{-1}(\Delta)=\xi(\Delta),$$

and we shall call ξ also a random variable.

This correspondence has the following properties:

(i) $\xi(\phi)=\phi\in\mathcal{S}$; $\xi(\mathbb{R})=\Omega$.

(ii) If $\Delta_i\perp\Delta_k$ for $i\neq k$ then $\xi(\Delta_i)\perp\xi(\Delta_k)$

(disjoint sets are mapped into disjoint sets).

(iii) $\xi(\bigcup \Delta_n)=\bigcup_n \xi(\Delta_n)$

for any pairwise disjoint sequence Δ_n.

If X_0, X_1, and X_2 are random variables, i.e., measurable functions, then so are X_1+X_2, X_1X_2, X^{-1} (if it exists) and for any sequence X_n ($n=1, 2,\ldots$) $\lim\sup X_n$, $\lim\inf X_n$, and $\lim X_n$ (if the limit exists).

4. *The Distribution Function*

Let X be a random variable and denote by

$$S_a\equiv\xi((-\infty, a])$$

then
$$F_\xi(a)=\mu(S_a)$$
is called the *distribution* function of the random variable ξ in the probability space $(\Omega, \mathscr{S}, \mu)$.

It has the following properties:

(i) $F_\xi(a)$ is a nondecreasing function, continuous from the right and it tends to the limit 0 as $a\to-\infty$.

(ii) $F_\xi(\infty)=1$.

If $F_\xi(a)=1$ is continuous and absolutely continuous then we may define a probability density $f_\xi(a)\geqslant$ by setting
$$\frac{\mathrm{d}F_\xi(a)}{\mathrm{d}a}=f_\xi(a).$$
The derivative exists everywhere.

5. *The Expectation Value*

Let ξ be a random variable, $F(a)$ its distribution function, then we define the *expectation* value by the integral (if it exists)
$$\langle\xi\rangle=\int_{-\infty}^{+\infty} a\,\mathrm{d}F(a)=E(X).$$
This is also called the *mean value* of ξ in $(\Omega, \mathscr{S}, \mu)$.

The notation is chosen deliberately in order to adumbrate the proposed generalization. The expression on the right-hand side is the classical one, while the left-hand side is used for the quantal one.

If ξ_1 and ξ_2 are two random variables represented by their measurable functions X_1 and X_2 we denote by $\xi_1+\xi_2$ the random variable represented by X_1+X_2. Similarly if ξ is represented by X then ξ^2 is represented by X^2.

With this notation we find for the *variance*
$$D^2(\xi)=\langle(\xi-\langle\xi\rangle)^2\rangle=E((X-E(X))^2)$$
or
$$D^2(\xi)=\langle\xi^2\rangle-\langle\xi\rangle^2=E(X^2)-E(X)^2.$$
The notion of independent random variable is of great importance in probability calculus. We formulate it here also in a generalizable fashion first for sets.

Two sets $A, B \in \mathscr{S}$ are said to be independent with respect to the probability measure μ if

$$\mu(A \cap B) = \mu(A)\, \mu(B).$$

The notion can be generalized to n sets $A_1, A_2, \ldots, A_n \in \mathscr{S}$. They are independent if and only if for every $i_1, i_2, \ldots, i_m$ $(m \leqslant n)$

$$\mu(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}) = \mu(A_{i_1})\, \mu(A_{i_2}) \ldots \mu(A_{i_m}).$$

The notion can be extended to random variables. The random variables ξ and η represented by the measurable functions X and Y are independent with respect to μ if for any pair $A, B \in \mathfrak{B}(\mathbb{R})$ of Borel sets on the real line

$$\mu(\xi(A) \cap \eta(B)) = \mu(\xi(A))\, \mu(\eta(B)).$$

These are the essential concepts of the classical probability calculus.

IV. THE PROBABILITY CALCULUS IN CLASSICAL MECHANICS

For a classical mechanical system the probability space Ω is the classical phase Γ. The probability measure for a system with no restriction will be the Lebesgue measure on Γ. Lionville's theorem assures that this measure is invariant under the evolution of the system due to the classical equations of motion.

Actually in isolated systems it is not this measure which can be used since it is not normalizable to one. Isolated systems will be restricted to a surface of constant energy. This measure is called the microcanonical measure and it is only defined on the surfaces of constant energy. If the system is not isolated but kept at a constant temperature by thermal contact with a heat bath then it is the canonical measure which is appropriate.

Every state of the system defines a new kind of measure. In particular a 'pure' state is given by a measure concentrated in one point $\omega \in \Gamma$. We shall denote it by δ_ω. It is defined explicitly by

$$\delta_\omega(A) = \begin{cases} 1 & \text{for} \quad \omega \in A \\ 0 & \text{for} \quad \omega \notin A. \end{cases}$$

The distribution function $F(a)$ of a random variable ξ for a pure state

δ_ω is defined by

$$F(a)=\delta_\omega(\xi((-\infty, a]))=\begin{cases}1 & \text{for} \quad \omega\in\xi((-\infty, a])\\ 0 & \text{for} \quad \omega\notin\xi((-\infty, a]).\end{cases}$$

For such a state the expectation value of the random variable ξ is given by

$$\langle\xi\rangle=\int_{-\infty}^{+\infty} a \, \mathrm{d}F(a)=a_0$$

where a_0 is the smallest value a for which $\xi((-\infty, a])=1$.

V. THE PROBABILITY CALCULUS IN QUANTUM MECHANICS

The preceding discussion of the probability calculus in classical mechanics serves the purpose of illustrating the need for generalizing this calculus if it is intended for application in quantum physics.

The first important observation is the absence of the phase space Γ in quantum mechanics. Hence it is necessary to develop a probability calculus without Kolmogorov's set Ω used for the definition of the measure space. At first sight this seems impossible since it would seem to make the definition of random variables impossible. However this is not so.

A careful examination of the classical probability calculus reveals that it could have been developed without ever mentioning the set Ω. The only place where this is not obvious is in the definition of random variables which we have defined as measurable functions $X(\omega)$ of $\omega\in\Omega$. However the subsequent use of these functions consisted merely in establishing σ-homomorphism $\xi:\mathfrak{B}(\mathbb{R})\to\mathcal{S}$ through the formula

$$\xi(\Delta)=X^{-1}(\Delta)\in\mathcal{S} \quad \text{for all} \quad \Delta\in\mathfrak{B}(\mathbb{R}).$$

Hence the calculus can be reconstructed in its entirety without ever mentioning Ω if we define random variables by this homomorphism. Of course in this case the class $\mathcal{S}$ must no longer be considered as consisting of the subsets of a set. Instead we replace it by a set of elements for which union, intersection, and complement is defined, in short $\mathcal{S}$ is a lattice.

In the classical case the lattice $\mathcal{S}$ was of a special kind, called a Boolean lattice, which is characterized by the distributive law

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \tag{0}$$

Once we have freed ourselves from the special interpretation of the lattice $\mathscr{S}$ as subsets of a set there is no need for maintaining the distributive law.

The structure of the lattice of 'elementary events' – we shall call them propositions, or yes-no experiments – will have to be determined from experiment and this involves a physical interpretation of the operations $\cap$, $\cup$ and the complement. This has been done in the case of quantum mechanics, and the result is that the operations of union and intersection lead to a non-Boolean lattice. This is the essential feature of general quantum mechanics.

I should perhaps mention here for completeness that there have been attempts to represent the quantum-mechanical proposition system on a weaker structure, the partially ordered sets (or posets).[9] The reason is that it is not always possible to exhibit in an operational manner the meet $A \cap B$ of two elementary events. However, as shown in Jauch (1968),[10] there are situations where this is possible even for a noncompatible pair of propositions A, B. This is always the case if there exist two passive filters, which represent measurements of the first kind corresponding to these two propositions. There exists then a filter $A \cap B$ which is obtained as an infinite alternating sequence of filters A and B. This is the operational analogue of the well-known formula $E \cap F = S - \lim_{n \to \infty} (EF)^n$ for the meet of two not necessarily commuting projection operators E and F in Hilbert space.

We shall denote by $\mathscr{L}$ the lattice of elementary events (propositions) in quantal physics and by $a, b, c, \ldots$ the elements from $\mathscr{L}$.

We have a partial-order relation in $\mathscr{L}$ denoted by c, as well as the operations of join and meet $a \cup b$ and $a \cap b$. They define the greatest lower bound and the least upper bound of a and b.

The lattice of propositions is orthocomplemented. The orthocomplement of a is denoted by a' and it satisfies

$$a \subset b \Rightarrow b' \subset a'$$
$$a \cap a' = \phi$$
$$a \cup a' = I,$$

where ϕ is the smallest and I the largest element in the lattice. These elements take the role of the null set and the entire set in the classical case.

Two propositions $a, b \in \mathscr{L}$ are said to be *disjoint* if $a \subset b'$ where b' is the orthocomplement of b. In this form the definition of disjointness is identical with the classical one. The notation for this relation is also $a \perp b$. The lattice has a smallest and a largest element denoted by ϕ and by I, respectively. In a sense the role of Ω in the classical case is taken now by the element $I \in \mathscr{L}$.

A probability measure on $\mathscr{L}$ is a function $\mu: \mathscr{L} \to [0, 1]$ defined on $\mathscr{L}$ with values in $[0, 1]$ satisfying the following conditions

(i) $\Sigma \mu(a_i) = \mu(\bigcup a_i)$ for $a_i \in \mathscr{L}$, $i = 1, 2, \ldots, a_i \perp a_k$ for $i \neq k$.

(ii) $\mu(\phi) = 0$, $\mu(I) = 1$.

(iii) If $\mu(a) = \mu(b) = 1$ then $\mu(a \cap b) = 1$.

The first two properties are exactly as in the classical calculus; the third is new. In fact in the classical calculus the third is a consequence of the other two. In the quantal calculus it is independent and therefore has to be postulated separately.

Passing now to the definition of random variables we use the definition which does not refer to the space Ω.

DEFINITION. *A* random variable *is a σ-homomorphism $\xi: \mathfrak{B}(\mathbb{R}) \to \mathscr{L}$ from the Borel sets on the real line into the lattice $\mathscr{L}$ of propositions, which satisfies the following conditions.*

(i) $\xi(\phi) = \phi$, $\xi(\mathbb{R}) = I$.

(ii) *For any disjoint sequence $\Delta_i \in \mathfrak{B}(\mathbb{R})$ ($\Delta_i \perp \Delta_k$,* for *$i \neq k$)*
$\xi(\bigcup \Delta_i) = \bigcup_i \xi(\Delta_i)$.

(iii) $\Delta_1 \perp \Delta_2 \Rightarrow \xi(\Delta_1) \perp \xi(\Delta_2)$.

An immediate consequence of these properties is that the range of the map ξ is a Boolean sublattice of $\mathscr{L}$. This is due to the fact the map is a homomorphism, that is, it conserves the lattice structure, which means that

$$\xi(\Delta_1 \cup \Delta_2) = \xi(\Delta_1) \cup \xi(\Delta_2)$$
$$\xi(\Delta_1 \cap \Delta_2) = \xi(\Delta_1) \cap \xi(\Delta_2)$$
$$\xi(\Delta') = \xi(\Delta)'.$$

From this follows that the image of the map is a Boolean sublattice of $\mathscr{L}$. The distribution function $F_\xi(a)$ is defined, as before, by

$$F_\xi(a)=\mu(\xi(S_a)) \quad \text{with} \quad S_a=(-\infty, a].$$

It induces a Stieltjes-Lebesgue measure μ_ξ on the Borel sets Δ.

The expectation value of a random variable ξ is then defined by

$$\langle\xi\rangle=\int_{-\infty}^{+\infty} a \,\mathrm{d}F_\xi(a).$$

From the foregoing we see that there is a close analogy between the classical and the quantal probability calculus. But there is also a profound difference due to the fact that the lattice of yes-no experiments for a quantal system is non-Boolean. The difference becomes explicit when we study the notion of *joint probability distribution* of two random variables.

It is useful to begin with the notion of compatibility. Two elements $a, b \in \mathscr{L}$ are said to be *compatible* and we denote this relation with $a \leftrightarrow b$ if the smallest sublattice which contains a, b, a', and b' is Boolean. We call this the lattice *generated* by a and b.

It is easy to see that a sublattice $\mathfrak{B} \subset \mathscr{L}$ is Boolean if and only if every pair of elements from $\mathfrak{B}$ is compatible.

The notion of compatibility can be transferred to random variables. To this end we define the ranges

$$\mathfrak{B}_\xi=\{a \mid a=\xi(\Delta), \Delta \in \mathfrak{B}(\mathbb{R})\}$$
$$\mathfrak{B}_\eta=\{a \mid a=\eta(\Delta), \Delta \in \mathfrak{B}(\mathbb{R})\}$$

and call ξ and η compatible if every $a \in \mathfrak{B}_\xi$ is compatible with every $b \in \mathfrak{B}_\eta$.

For pairs of classical random variables one can define the notion of *joint distribution*. It is defined as follows: let ξ and η be two classical random variables. The joint distribution is a function of two real variables a and b,

$$F_{\xi,\eta}(a, b)=\mu(\xi(S_a) \cap \eta(S_b)).$$

It is a nondecreasing function of both arguments satisfying the further conditions:

(i) $$F_{\xi,\eta}(-\infty, b)=F_{\xi,\eta}(a, -\infty)=0.$$

(ii) $$\int_{-\infty}^{+\infty} d_b F_{\xi,\eta}(a, b)=F_\xi(a)=\mu(\xi(S_a)).$$

(iii) $$\int_{-\infty}^{+\infty} d_a F_{\xi,\eta}(a, b)=F_\eta(b)=\mu(\eta(S_b)).$$

Since compatible random variables in the quantum probability calculus behave exactly like classical ones it is immediately obvious that such variables also have a joint probability distribution given by formulas identical with the preceding ones.

VI. RANDOM VARIABLES IN HILBERT SPACE

Before discussing the question of the distribution function of noncompatible random variables in quantum probability calculus we give the interpretation of random variables in Hilbert space.

It is known that every proposition system $\mathscr{L}$ admits a representation in a linear vector space with coefficients from the real, complex, or quaternion fields. This representation is particularly simple if the lattice is irreducible, or in physical terms, if the system admits no superselection rules.

The mathematical expression for this property is that the center $\mathscr{C}$ of $\mathscr{L}$ is trivial. With the *center* $\mathscr{C}$ we denote the set of elements which are compatible with every other element:

$$\mathscr{C}=\{a \mid a\in\mathscr{L},\ a\leftrightarrow x,\ \forall x\in\mathscr{L}\}.$$

Evidently $\phi\in\mathscr{C}$ and $I\in\mathscr{C}$. If these are the only two elements contained in $\mathscr{C}$ then we refer to $\mathscr{C}$ as being trivial.

The subspaces (or the projection operators) in a Hilbert space $\mathscr{H}$ form a lattice with $\phi=0$ (=zero projection), $I=I$ (=unit operator) $E'=I-E$ (orthocomplement), and $I\cap F=S-\lim_{n\to\infty}(EF)^n$ (=meet). The join is then defined by $E\cup F=(E'\cap F')'$. Under some mild additional restrictions one can show that the coefficients of the Hilbert space are the complex number field C.[11] We shall assume that this is the case. Under these hypotheses the abstract lattice of propositions is isomorphic to the lattice of subspaces of a Hilbert space $\mathscr{H}$, as it was demonstrated by Piron (1964).

Let us now examine what becomes of a probability measure and random variables in this case.

Let E_i $(i=1,2,\ldots)$ be a sequence of pairwise disjoint projections $(E_i \perp E_k$ or equivalently $E_i E_k = 0$ for $i \neq k)$, then a probability measure is a functional μ from the set of all projections $\mathscr{P}$ to the interval $[0, 1]$

$$\mu : \mathscr{P} \to [0, 1]$$

satisfying the three characteristic properties

(i) $\bigcup_i \mu(E_i) = \mu(\sum E_i)$.

(ii) $\mu(\phi) = 0, \qquad \mu(I) = 1$.

(iii) $\mu(E) = \mu(F) = 1 \Rightarrow \mu\,(E \cap F) = 1$.

According to a theorem due to Gleason,[12] if $\dim \mathscr{H} \geqslant 3$ every such measure can be represented by a positive trace class operator ρ of trace 1, such that

$$\mu(E) = \operatorname{Tr} \rho E.$$

In the special case that ρ is a projection operator of rank 1 we have $\rho^2 = \rho$ and if φ is in the range of ρ, so that $\rho\varphi = \varphi$, one obtains

$$\mu(E) = (\varphi, E\varphi).$$

In this manner we recover the usual expectation values for pure states as they occur in quantum mechanics.

Let us now consider a random variable in this setting. According to the definition of Part V, a real random variable is a σ-homomorphism $\xi : \mathfrak{B}(\mathbb{R}) \to \mathscr{P}$ from the Borel sets on the real line to the projections in $\mathscr{H}$, which satisfies the three conditions (i), (ii), and (iii) given in Part V.

An inspection of these conditions shows that these are exactly the conditions for the definition of a *spectral measure.* According to the spectral theorem every spectral measure defines uniquely a self-adjoint operator X according to the formula

$$X = \int_{-\infty}^{+\infty} \lambda \, \mathrm{d}E_\lambda$$

with

$$E_\lambda = \xi((-\infty, \lambda]) \equiv \xi(S_\lambda).$$

From Gleason's theorem follows then that the expectation value of ξ in the state μ is given by

$$\langle \xi \rangle = \mathrm{Tr}\, \rho X = \int_{-\infty}^{+\infty} \lambda \, \mathrm{d}\, \mathrm{Tr}(\rho E_\lambda).$$

Thus we have recovered all the usual formulas of quantum theory in Hilbert space.

I add a few comments to this result.

(1) I stated that property (iii) of the probability measure must be postulated since it cannot be derived from the other two as in the classical probability calculus. In order to appreciate this remark, I sketch the derivation of (iii) from the other two conditions in the classical case when $\mathscr{L}$ is a Boolean algebra.

THEOREM 1. *If $\mathscr{L}$ is a Boolean algebra, and μ is a function μ: satisfying conditions* (i) *and* (ii) *then*

$$\mu(a) = \mu(b) = 1 \Rightarrow \mu(a \cap b) = 1 \ \forall a, b \in \mathscr{L}.$$

Proof. If $a \cap b = \phi$ then they are disjoint. Hence by (i) $\mu(a) + \mu(b) = \mu(a \cup b) = 1$. Therefore $\mu(a) = 1 \Rightarrow \mu(b) = 0$ and $\mu(b) = 1 \Rightarrow \mu(a) = 0$. The hypotheses of the theorem cannot be satisfied.

We may thus assume that $a \cap b = c \neq \phi$. We may then write

$$a = a_1 \cup c$$
$$b = b_1 \cup c$$

where $a_1 = c' \cap a$, $b_1 = c' \cap b$, and a_1, b_1, c are pairwise disjoint. Hence from (i) we obtain

$$1 = \mu(a) = \mu(a_1 \cup c) = \mu(a_1) + \mu(c)$$
$$1 = \mu(b) = \mu(b_1 \cup c) = \mu(b_1) + \mu(c).$$

By taking the difference of these two questions we find first that

$$\mu(a_1)=\mu(b_1)\equiv x.$$

On the other hand from the sum of the two questions we obtain

$$(1) \qquad 1=x+\mu(c)$$

since $1=\mu(a)\leqslant\mu(a\cup b)\leqslant 1$ we have $\mu(a\cup b)=1$ and therefore from (i)

$$1=\mu(a\cup b)=\mu(a_1)+\mu(b_1)+\mu(c)$$

or

$$(2) \qquad 1=2x+\mu(c).$$

Comparing (1) with (2) we conclude that $x=0$ and therefore

$$\mu(a\cap b)=1 \quad \|.$$

(2) In the Hilbert space setting property (iii) can actually also be proved as a consequence of (i) and (ii) provided $\dim\mathscr{H}\geqslant 3$.

This is due to the following facts:

(a) Under this hypothesis every probability measure μ is of the form $\mu(E)=\operatorname{Tr}\rho E$ with ρ a positive trace class operator with trace 1;

(b) If E, F are any two projections then

$$E\cap F=S-\lim_{n\to\infty}(EF)^n;$$

(c) If T_n is a uniformly bounded sequence of operators and $T_n\to T$ strongly, then $\operatorname{Tr}\rho T$ exists and

$$\operatorname{Tr}\rho T_n\to\operatorname{Tr}\rho T,$$

where (a) is essentially Gleason's theorem quoted in this part, (b) is a well-known result on projections in Hilbert space (cf. note 1), (c) can be proved as follows: the operator ρ being of trace class may be written as

$$\rho=\sum_{r=1}^{\infty}\alpha_r P_r$$

where P_r are orthogonal projections which we may assume without loss of generality to be of rank 1.

The eigenvalues α_r may be ordered as a decreasing sequence

$$\alpha_1\geqslant\alpha_2\geqslant\cdots\geqslant\alpha_r\geqslant\cdots\geqslant 0.$$

Furthermore the trace condition means $\sum \alpha_r = 1$. Let $R < \infty$ be an integer such that $\sum_{R+1}^{\infty} \alpha_r < \varepsilon$ for some arbitrary $\varepsilon > 0$, and let $P_r \varphi_r = \varphi_r$, $\|\varphi_r\| = 1$. We obtain then for

$$|\mu(T_n) - \mu(T)| = |\sum \alpha_r (\varphi_r, (T_n - T)\varphi_r)|$$
$$\leqslant \sum_{r=1}^{R} \alpha_r |(\varphi_r, (T_n - T)\varphi_r)| + \sum_{R+1}^{\infty} \alpha_r |(\varphi_r (T_n - T)\varphi_r)|.$$

We now choose N such that for $n > N$

$$|(\varphi_r, (T_n - T)\varphi_r)| < \varepsilon \quad \forall (r = 1, 2, \ldots, R).$$

This is possible because $T_n \to T$ strongly, hence weakly. The first term becomes therefore

$$< \varepsilon \sum_{r=1}^{R} \alpha_r < \varepsilon.$$

For the second term we note that because $T_n \to T$ and T_n are uniformly bounded, T is also bounded, hence $|(\varphi_r, (T_n - T)\varphi_r)| \leqslant \|T_n\| + \|T\|$, so that the second term is

$$\leqslant \left(\sum_{R+1}^{\infty} \alpha_r\right)(\|\mathrm{T}_n\| + \|\mathrm{T}\|) \leqslant \varepsilon(\|\mathrm{T}_n\| + \|\mathrm{T}\|).$$

Because of the uniform boundedness the right-hand side is independent of n. Hence we have shown

$$\operatorname{Tr} \rho T \quad \text{exists an} \quad \operatorname{Tr} \rho T = \lim_{n \to \infty} \operatorname{Tr} \rho T_n.$$

Let us now verify property (iii). We note first that

$$\operatorname{Tr} \rho E = \sum_{r=1}^{\infty} \alpha_r (\varphi_r, E\varphi_r) = 1$$

implies

$$(\varphi_r, E\varphi_r) = \|E\varphi_r\|^2 = 1 \quad (r = 1, 2, \ldots, \infty),$$

so that

$$\|\varphi\|^2 = 1 = \|E\varphi\|^2 + \|(I - E)\varphi\|^2$$

or

$$\|(I - E)\varphi\|^2 = 0, \quad \text{or finally} \quad E\varphi = \varphi.$$

Thus for all r such that $\alpha_r > 0$

$$E\varphi_r = \varphi_r.$$

Similarly

$$F\varphi_r = \varphi_r.$$

Therefore

$$\sum \alpha_r(\varphi_r, EF\varphi_r) = \sum \alpha_r = 1,$$

so that

$$\operatorname{Tr} \rho EF = 1.$$

we denote $EF = T$, we note that $\|T^n\| \leqslant 1$ and conclude from the preceding reasoning that

$$\operatorname{Tr} \rho T^n = 1 \quad (n = 1, 2, \ldots).$$

Hence by (c)

$$\mu(E \cap F) = \operatorname{Tr} \rho E \cap F = 1 \quad \|.$$

(3) Property (iii) has a simple physical interpretation in case there exist passive filters corresponding to the propositions a and b. Indeed $\mu(a) = 1$ says that the filter corresponding to a is 100 percent transparent. Similarly $\mu(b) = 1$ implies that the filter corresponding to b is also 100 percent transparent. Since the filters are passive the system traverses the filters without modification of the state. Hence it will also traverse an infinite (or very large) alternating sequence of filters a and b. But such a sequence represents the filter corresponding to $a \cap b$. Hence $\mu(a \cap b) = 1$.

The only example known to me of a probability measure on a lattice which does not satisfy (iii) is in a lattice with a maximal chain of three elements. This is of course precisely the case that is excluded by the hypothesis of Gleason's theorem that $\dim \mathscr{H} \geqslant 3$. In view of this fact it would be of considerable interest to prove property (iii) in the lattice-theoretic setting. No such proof is known to me.

(4) The present derivation of Hilbert-space quantum theory from the lattice-theoretic one elucidates the relation between *compatibility* of observables and *commutativity* of the corresponding operators in Hilbert space.

The former is a physical property and the latter a mathematical one. In the light of the phenomenological interpretation of the lattice structure, compatibility is represented by the relation $a \leftrightarrow b$ which in turn is equivalent to the property that the sublattice generated by (a, b, a', b') is Boolean. This is exactly how it is in classical physics, where every such sublattice is Boolean since the entire proposition system $\mathscr{L}$ is.

In the representation of the proposition system by the subspaces of a Hilbert space, compatibility of two projection operations E, F is equivalent to commutability of these operators. We have in fact the following.

THEOREM 2. *$\mathscr{L}(E, F, E', F')$ is Boolean* $\Leftrightarrow [E, F]=0$.

Proof. If $\mathscr{L}(E, F, E', F')$ is Boolean then

$$E=E_1+G$$
$$F=F_1+G$$

with

$$G=E \cap F, \qquad E_1=E \cap G', \qquad F_1=F \cap G'.$$

It follows that

$$EF=(E_1+G)(F_1+G)=G$$
$$FE=(F_1+G)(E_1+G)=G.$$

Therefore

$$[E, F]=0 \quad \|.$$

If $[E, F]=0$, then $E \cap F = EF$. Hence for any triplet, for instance E, F, F', we have

$$E \cap (F \cup E')=(E \cap F) \cup (E \cap E')=(E \cap F) \cup \phi = EF.$$

But

$$F \cup E' = I - E + FE,$$

so that

$$E \cap (F \cup E')=E(I-E+FE)=EFE=FE=EF.$$

Thus for any triplet chosen from E, F, E', F' we have the distributive law and this implies that $\mathscr{L}(E, F, E', F')$ is Boolean. $\|$

This result disagrees with the opinion expressed by Park and Margenau in a recent publication (see Park and Margenau, 1968). However it is seen that this result is independent of the hypothesis whether the correspondence between observables and self-adjoint operators is one to one contrary to what is claimed in that reference. In fact the essential hypothesis is the much weaker one that with the propositions $a, b \in \mathscr{L}$ the proposition 'a and b' $= a \cap b$ is also contained in $\mathscr{L}$.

VII. JOINT DISTRIBUTIONS OF RANDOM VARIABLES

In the new quantum probability calculus there is an essential feature which distinguishes it from the classical calculus. This is the occurrence of noncompatible observables or random variables. In the classical calculus every observable is compatible with every other one, due to the fact that the lattice $\mathscr{L}$ is Boolean. In the quantum calculus this is not necessarily the case.

In the classical case it was possible to define the joint distribution function of two random variables ξ, η as the function $F_{\xi,\eta}(a, b)$ satisfying the following properties

(0) $F_{\xi,\eta}(a, b) \geqslant 0$ and nondecreasing in a and b.

(1) $F_{\xi,\eta}(-\infty, b) = F_{\xi,\eta}(a, -\infty) = 0$.

(2) $$\int_{-\infty}^{+\infty} d_b F_{\xi,\eta}(a, b) = F_\xi(a).$$

(3) $$\int_{-\infty}^{+\infty} d_a F_{\xi,\eta}(a, b) = F_\eta(b).$$

In the quantal case the definition of such a joint probability may be impossible in case the random variables ξ and η are not compatible. This corresponds to the physical fact that joint measurements of arbitrary noncompatible variables may be impossible.

Since the preceding statement is flatly contradicted by Park and Margenau,[13] I must interpose at this point a few critical remarks concerning their analysis of the measuring process in quantum theory.

Their analysis concerns primarily the notion of pairs of incompatible observables. They insist that in spite of the uncertainty relation, such as $\Delta p \Delta q \geqslant \frac{1}{2}\hbar$ for canonical variables p and q, such variables can be measured with arbitrary degree of accuracy. They therefore reject complementarity, so essential in Bohr's analysis of the quantal systems, and they believe this concept can be replaced by the simpler notion of 'latency'. Although they would agree that the uncertainty relation is valid for measurements on an ensemble of identically prepared systems, they believe that this relation is not a restriction concerning the accuracy of measurements for complementary variables of an individual system.

The essential point in their analysis concerns the 'joint' measurements of noncommuting observables. Although an explicit definition of their notion of compatibility is never given in their paper one gathers from the context that for them compatibility means that a joint measurement of the pair of observables is possible. By showing that certain pairs of noncommuting observables are measurable simultaneously to an arbitrary degree of accuracy they come to the conclusion that noncommuting observables may very well be compatible in their sense of the term. (Incidentally it is not clear from their paper whether they believe that any pair of noncommuting observables is compatible in this sense or not.)

They conclude from this that there are joint measurements possible for certain variables such as p and q even though neither a joint probability distribution nor an operator exists for representing such joint measurements.

In order to appraise this point of view it is necessary to recall that in Bohr's point of view the arbitrary precision of individual measurements of canonical variables such as p and q was never in question. Both quantities can in principle be measured with a precision only limited by the inherent precision of the applied experimental arrangement. However the very presence of this experimental arrangement precludes the simultaneous attributions of precisions to complementary variables, such as p and q which would violate the uncertainty relation.

The example given by Park and Margenau for such a measurement is no counterexample to this general and essential feature of quantal systems. Their example is in fact only a determination of the position q with a given accuracy Δq followed by a determination of p a long time t later with an arbitrary accuracy Δp. Their conclusion that this second measurement permits them to assert that this also constitutes a measurement of p with that same accuracy at time $t=0$ is not correct. Their justification for this is that the probability distribution of p at time $t=0$ is the same as at a time $t>0$. While this statement is perfectly correct it is not sufficient for asserting that the actual value of the p at the two times is equal.

With this counterexample shown to be irrelevant for the question under discussion their entire case falls to the ground and the difficulties which they had to face concerning joint distributions of noncommuting observables disappear.

Returning now to the problem of the joint probability distribution for incompatible observables it is very easy to see in the quantum probability calculus that such a distribution cannot exist satisfying the properties listed above for the canonical variables.

The reason for this is the fundamental relation

$$\xi(S_a) \cap \eta(S_b) = \phi \quad \text{for} \quad -\infty < a, b < +\infty. \tag{1}$$

Indeed if $\xi(S_a) \cap \eta(S_b)$ were $\neq \phi$ then there would exist a function $\varphi(x) \in L^2(-\infty, +\infty)$ with the properties

$$\varphi(x) = 0 \quad \text{for} \quad x > a$$
$$\hat{\varphi}(x) = 0 \quad \text{for} \quad x > b$$

where $\hat{\varphi}$ is the Fourier transform of φ. It is well known that such a function does not exist unless $\|\varphi\| = 0$.

Due to the relation (1) it follows that

$$F_{\xi,\eta}(a, b) = 0 \quad \text{for} \quad -\infty < a, b < +\infty$$

so that properties (2) and (3) are violated.

But this negative conclusion does not preclude that p and q (or in fact any pair of noncommuting observables) are measurable within an accuracy limited by the uncertainty relation. Thus a joint probability distribution should exist in a more restrictive sense which is in accord with this restriction.

In order to define this weaker sense we modify the definition of $F_{\xi,\eta}$. Instead of a nondecreasing function on $\mathbb{R}^2$ we define it as a finitely additive set function on the Borel rectangles. If ξ and η are compatible then this definition is possible and the function $F_{\xi,\eta}$ satisfies the properties

$$F_{\xi,\eta}(A \times B) \geq 0 \tag{0}$$

(1) $F_{\xi,\eta}(\phi \times B) = F_{\xi,\eta}(A \times \phi) = 0$
(2) $F_{\xi,\eta}(A \times \mathbb{R}) = F_{\xi}(A) = \mu(\xi(A))$
(3) $F_{\xi,\eta}(\mathbb{R} \times B) = F_{\eta}(B) = \mu(\eta(B))$.

According to a well-known theorem in measure theory such a finitely additive set function on Borel rectangles has a unique extension to the Borel sets on $\mathbb{R}^2$, defining a product measure on $\mathbb{R}^2$.

For noncompatible random variables, such as p and q, it is still possible to define a function $F_{\xi,\eta}$ satisfying all the properties listed above by setting

$$F_{\xi,\eta}(A \times B) = \mu(\xi(A) \cap \eta(B)).$$

But in agreement with Wigner's (1932) result this function is not an additive set function on Borel rectangles and therefore cannot be extended to a measure on $\mathbb{R}^2$.

In spite of this anomaly the function $F_{\xi,\eta}(A \times b)$ is not entirely devoid of physical meaning. It represents in fact the probability that in a given state the variable ξ assumes values in the set A while at the same time the variable η assumes values in the set B.

This probability is not necessarily zero as may be seen in the case $\xi = p$ and $\eta = q$. In this case as we have noted before $\xi(A) \cap \eta(B) = \phi$ if $m(A')\, m(B') = \infty$.

But in case $m(A')\, m(B') < \infty$ this is not true. We have in fact the following.

THEOREM 3. *If $E \equiv E_A$, $F \equiv F_B$ represent the spectral projections of the canonical variables p and q associated respectively with the Borel sets A and B, then*

$$m(A')\, m(B') < \infty \Rightarrow E \cap F \neq \phi.$$

The proof of this theorem will be given elsewhere. Suffice it here to remark that the theorem implies the following statement concerning functions $\varphi \in L^2(-\infty, +\infty)$ and their Fourier transform $\hat{\varphi}$. We shall say that φ has a gap (of positive measure) if there exists a Borel set A with Lebesgue measure $m(A)$ such that

$$0 < m(A') < \infty.$$

The theorem then asserts that there exist functions $\varphi(x)$ with a gap whose Fourier transform also has a gap.[14]

In conclusion I may thus state that, although joint measurements of certain noncompatible observable may be possible with nontrivial results, it is not true that there exists an observable which represents all the joint measurements of two such observables.

University of Geneva

NOTES

* This article was written during a visit to the University of Colorado in Boulder. It is a pleasure to thank Professor K. Gustafson, who made this visit possible. Thanks are also due to Professor B. Misra for an important remark concerning the joint distribution of noncompatible observables.

[1] The quantum probability calculus was briefly sketched by Jauch (1968). The role of probability in quantum theory was the principal subject of three papers by Suppes (1961, 1963, 1966).

[2] For a detailed review of hidden variable theories the reader is referred to Belinfante (1973).

[3] This point was first made by Wigner (1932). It was the subject of many subsequent papers such as Moyal (1949), Brittin and Chappell (1962), Park and Margenau (1968).

[4] For recent papers on this subject I refer to Gudder (1967, 1968), Gudder and Marchand (1972), and Varadarajan (1962).

[5] The principal difference with respect to some other work in this field is that the quantal proposition system is assumed to be a lattice and not just an orthocomplemented partially ordered set (poset). The empirical justifications for this assumption were first given by Piron (1964), where it was shown quite explicitly that for many physical systems the poset structure is not sufficient for representing the phenomenology. Further details are given by Jauch (1968).

[6] This useful distinction is due to my late friend, Dr. G. Baron, whose profound knowledge of fundamental problems on probability theory has greatly influenced my thinking on the subject.

[7] Kolmogorov (1956). This is the English version of the original German version.

[8] Renyi (1970) introduced the probability calculus based on the basic notion of relative probability. This generalizes Kolmogorov's calculus to nonnormalizable probability fields.

[9] This form of quantum probability calculus was first developed by Varadarajan (1962).

[10] Jauch (1968) uses for instance the construction of composite filters. There are other possibilities of constructing the meet of two elementary noncompatible events.

[11] The question of the number field remained for a long time beyond an empirical test. The recent work by Gudder and Piron (1971) is the best that one can do.

[12] The conjecture that every σ-additive measure on orthogonal subspaces is given by a density matrix was finally proved by Gleason (1957). Similar conjectures on the projection lattice of von Neumann algebras remain unproved.

[13] Park and Margenau (1968) claim to have shown that measurements are possible which violate the uncertainty relations.

[14] I am indebted to Prof. Martin Peter for an explicit construction of such functions,

who also showed that their existence is not without physical interest especially in the theory of metals. The question whether such functions exist and their relevance for the problem of joint distributions was first mentioned to me by Prof. A. Galindo of Madrid.

REFERENCES

Belinfante, F. J., *A Survey of Hidden Variable Theories*, Pergamon Press, New York, 1973.

Brittin, W. E. and Chappell, W. R., 'The Wigner Distribution Function and Second Quantization in Phase Space', *Review of Modern Physics* **34** (1962), 620–627.

Gleason, A. M., 'Measures on the Closed Subspaces of a Hilbert Space', *Journal of Mathematics and Mechanics* **6** (1957), 885–893.

Gudder, S. P., 'Hilbert Space, Independence, and Generalized Probability', *Journal of Mathematical Analysis and Applications* **20** (1967), 48–61.

Gudder, S. P., 'Joint Distributions of Observables', *Journal of Mathematics and Mechanics* **18** (1968), 325–335.

Gudder, S. P. and Marchand, J. P., 'Non-Commutative Probability on von Neumann Algebras', *Journal of Mathematical Physics* **13** (1972), 799–806.

Gudder, S. P. and Piron, C., 'Observables and the Field in Quantum Mechanics', *Journal of Mathematical Physics* **12** (1971), 1583–1588.

Jauch, J. M., *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, 1968.

Kolmogorov, A. N., *Foundations of the Theory of Probability*, Chelsea, New York, 1956, transl. and ed. by N. Morrison.

Moyal, J. E., 'Quantum Mechanics as a Statistical Theory', *Proceedings of the Cambridge Philosophical Society* **45** (1949), 99–124.

Park, J. L. and Margenau, H., 'Simultaneous Measurability in Quantum Theory', *International Journal of Theoretical Physics* **1** (1968), 211–282.

Piron, C., 'Axiomatique Quantique', *Helvetica Physica Acta* **37** (1964), 439–467.

Renyi, A., *Probability Theory*, North Holland, Amsterdam, 1970.

Suppes, P., 'Probability Concepts in Quantum Mechanics', *Philosophy of Science* **28** (1961), 378–389.

Suppes, P., 'The Role of Probability in Quantum Mechanics' in B. Baumrin (ed.), *Philosophy of Science, The Delaware Seminar*, Wiley, New York, 1963.

Suppes, P., 'The Probabilistic Argument for a Non-classical Logic of Quantum Mechanics', *Philosophy of Science* **33** (1966), 14–21.

Varadarajan, V. S., 'Probability in Physics and a Theorem on Simultaneous Observability', *Comments on Pure and Applied Mathematics* **15** (1962), 189–216.

Wigner, E., 'On the Quantum Theory for Thermodynamic Equilibrium', *Physical Review* **40** (1932), 749–759.