Foundations of Physics, Vol. 22, No. 11, 1992

On the Structure of the Quantum-Mechanical Probability Models[1]

Nicola Cufaro-Petroni[2]

Received February 12, 1992

In this paper the role of the mathematical probability models in the classical and quantum physics is shortly analyzed. In particular the formal structure of the quantum probability spaces (QPS) is contrasted with the usual Kolmogorovian models of probability by putting in evidence the connections between this structure and the fundamental principles of the quantum mechanics. The fact that there is no unique Kolmogorovian model reproducing a QPS is recognized as one of the main reasons of the paradoxical behaviors pointed out in the quantum theory from its early days.

1. INTRODUCTION

In a statistical theory it is important to make a clear difference between the sets of empirical data and the mathematical (probabilistic) model built in order to describe them. This remark points out, on the one hand, the fact that, in general, there are many possible models for the same set of experimental results and, on the other, the fact that (being a mathematical model only a useful abstraction) nothing compels us to strictly adhere to one particular formal structure of these models. This remark is particularily suitable for a discussion on quantum mechanics since the difference between the classical and the quantum physics (as enforced on us by the experimental results and described by means of general ideas such as the superposition and the uncertainty principles) can in some sense be described as the difference of the mathematical models that can be adapted

[1] Paper written in honor of L. de Broglie.

[2] Dipartimento di Fisica dell' Universitá ed Istituto Nazionale di Fisica Nucleare, via Amendola 173, 70126 Bari, Italy.

0015-9018/92/1100-1379$06.50/0

to a classical or to a quantum experiment. The aim of this paper is to discuss some aspects of this situation by reviewing the peculiar probabilistic behaviors induced by the fundamental principles of quantum mechanics along with the similarities between the classical and quantum probabilistic models. We think in fact that in this direction the deep intuitions of L. de Broglie about wave functions, trajectories of quantum particles, corpuscular interference, and so on have not exhausted their impressive power of suggestion and that a lot of stuff is still there to be understood. In this sense this article can be considered as a contribution to the still running discussion about the real meaning of the quantum theory and a step in a better comprehension of its implications.

We start with a short qualitative discussion of what a statistical theory can be. Let X, Y,... indicate the measurable quantities of our physical world, and let us suppose that they take the real values x, y,.... In correspondence with a given *preparation* (namely a measurement of some physical quantities, followed by a selection of the results) $\mathscr{U}$ of the system, we are able to calculate probabilities like $\Pr(X \in B \backslash \mathscr{U})$, $B \subseteq \mathbf{R}$ by means of a suitable frequentistic approach. Here the idea of conditioning is in fact already contained in the previous definition. Conditioning means acquisition of information in order to evaluate the likelihood (probability) of particular events. In this sense even our previous definition is a conditional (through $\mathscr{U}$) probability. However, in the usual language the name of conditional probability is reserved for subsequent (to $\mathscr{U}$) conditionings. To be precise, we should say that, if we add to our previous definition the information derived from another, subsequent preparation $\mathscr{V}$, our probabilities will be, in general, modified: $\Pr(X \in B \mid \mathscr{U}, \mathscr{V})$. We remark here that no precise relation between $\Pr(X \in B \mid \mathscr{U})$ and $\Pr(X \in B \mid \mathscr{U}, \mathscr{V})$ is given *a priori*. Of course we can also define average values of our quantities $\mathbf{M}(X \mid \mathscr{U})$, new quantities as functions of other quantities: $Z = f(X)$, and all sorts of statistical objects.

We claim that we have a classical (Kolmogorovian) probabilistic model of our set of statistical data when we can build a probability space $(\Omega, \mathscr{F}, \mu)$ (Ω is a set, $\mathscr{F}$ a σ-algebra of parts of Ω, and μ a probability measure on $\mathscr{F}$) so that the physical quantities X, Y,... can be represented by measurable functions (random variables) ξ, η,... and that there is a correspondence between measured and calculated probabilities: If μ corresponds to the preparation $\mathscr{U}$, we require $\Pr(X \in B \mid \mathscr{U}) = \mu(\xi \in B)$. In particular, if $\mathscr{V}$ is a subsequent preparation corresponding to the verification of the event $(Y \in A)$, $A \subseteq \mathbf{R}$, we require that

$$\Pr(X \in B \mid \mathscr{U}, Y \in A) = \frac{\mu(X \in B, Y \in A)}{\mu(Y \in A)}$$

This rule connecting *a priori* probabilities and conditional probabilities is in fact peculiar of the Kolmogorovian models and in no way can be extended to every probabilistic model. Moreover, it should be remarked here that in this relation the idea of *subsequent* preparations, which is present on the left-hand side, is replaced on the right-hand side by the idea *simultaneous* events which is present in the joint statement. We point out this because, as we will see later, quantum models are characterized by the fact that joint statements are no longer always possible, so that it is very well understandable that the previous relation no longer holds. Of course, in this classical model there is a formal correspondent for every statistical object; for example, the Lebesgue integrals (expectation values) correspond to the mean values: $\mathbf{E}(X) = \mathbf{M}(X \mid \mathscr{U})$. In particular, it is well known that an important role is played by indicators of subsets $D \subseteq \Omega$:

$$I_D(\omega) = \begin{cases} 1, & \omega \in D \\ 0, & \omega \in \bar{D} \end{cases}$$

(where $\bar{D}$ is the complement of D). For example, we could describe our Kolmogorovian model in terms of algebras of indicators (instead of events), and expectation values (instead of probabilities). In this case unions, intersections, complements,... of subsets of Ω are replaced by sums, differences, products,... of indicators; and we have $\mu(D) = \mathbf{E}(I_D)$.

In the next section we will try to reproduce in the quantum case the previous well-known classical construction of the basic elements of a Kolmogorovian model, and, in order to explore similarities and differences, we will consider the quantum mechanics as a well-defined formal theory[3] (we will refrain, in this paper, from criticizing or modifying it), and we will take for granted that quantum physics can be correctly described in a Hilbert space H where the physical quantities X, Y,... are represented by self-adjoint operators $\hat{X}$, $\hat{Y}$,... (corresponding to classical random variables) and statements are represented by projection operators $\hat{P}_L$, $\hat{P}_M$,... (corresponding to indicators of classical subsets of Ω) on closed subspaces L, M,.... Moreover, we will always limit our consideration to the case of a finite number of statements in order to avoid the discussion of convergence problems, which is not our principal aim here.

2. QUANTUM PROBABILITY MODELS

Let us introduce the following notations: given a Hilbert space H with vectors x, y,... and scalar product (x, y), we will indicate with L, M,... its

[3] See, for example, Refs. 1 or 2.

closed linear subspaces and with $\hat{P}_L$, $\hat{P}_M$,... the corresponding projections. To emphasize that a closed subspace is spanned by the vectors x, y,... we will use the notation $[x, y,...]$. Moreover: LM will be the closed subspace consisting of the vectors belonging to both L and M; $M+N$ will be the subspace consisting of all the linear combinations of vectors in L and M; finally $\bar{L}$ is the subspace of all the vectors orthogonal to L. The proof of the following relations is immediate:

$$\overline{L+M}=\bar{L}\bar{M}$$

$$\bar{\bar{L}}=L$$

$$\overline{LM}=\bar{L}+\bar{M}$$

Furthermore, given two closed subspaces L and M, we will indicate with $L-M$ the closed subspace of the vectors of L which are orthogonal to LM. Finally we will indicate with $[0]$ the subspace of H containing only the null vector 0 and with $\hat{I}$, $\hat{0}$ respectively the identity and the null operators.

Definitions 1. Two subspaces L, M are said to be *orthogonal*, and we indicate it with $L \perp M$, when all the vectors of L are orthogonal to all the vectors of M.

Remark 1. When $M \subseteq L$, then $L-M$ is the subspace of the vectors of L orthogonal to M, namely, $L-M=L\bar{M}$; when $LM=[0]$, then $L-M=L$. Furthermore, we always have that $L-M=L\overline{LM}$ and $\bar{L}=H-L$.

Proposition 1. $L-M \perp M \Leftrightarrow M-L \perp L$.

Proof. Suppose that $L-M \perp M$ and take $x \in M-L$ (namely, $x \in M$ and $x \in \overline{LM}$) and $y \in L$. Since LM is a closed subspace of L, there is one and only one decomposition of y such that

$$y=u+v \qquad u \in LM, \qquad v \in L-M$$

Hence, since $u \in LM$ and $x \perp LM$, we have

$$(x, y)=(x, u)+(x, v)=(x, v)$$

But $v \in L-M$ and $x \in M$, so that, by hypothesis, they are orthogonal and hence $(x, y)=0$. The inverse is proven in exactly the same way. □

Definition 2. We will say that L and M are *weakly orthogonal* or *w-orthogonal*, and we indicate it with $\oplus M$, when $L-M \perp M$ or equivalently (see Proposition 1) when $M-L \perp L$.

Remark 2. When two subspaces are orthogonal, they are also w-orthogonal, but the inverse does not hold: this is the rationale for this new notion. For example, in a three-dimensional Euclidean space endowed with three orthogonal axis, two coordinate planes are w-orthogonal but not orthogonal. However, if $LM = [0]$, the notions of orthogonality and of w-orthogonality become equivalent. Of course, as a consequence, $L \obot \bar{L}$ always holds.

Remark 3. The binary relation $\obot$ is symmetric (see Proposition 1) and reflexive:

$$M - M = M\overline{MM} = [0] \perp M \Rightarrow M \obot M$$

but it is not transitive, as can be seen by remarking that even the usual orthogonality is not transitive.

Proposition 2. $[\hat{P}_L, \hat{P}_M] = \hat{0} \Leftrightarrow L \obot M$.

Proof. Let us first suppose that $[\hat{P}_L, \hat{P}_M] = \hat{0}$; it is known[4] that this implies $\hat{P}_L \hat{P}_M = \hat{P}_{LM}$. If now we take $x \in L - M$ (namely, $x \in L$, $x \perp LM$) and $y \in M$, we will have

$$\hat{P}_L x = x, \qquad \hat{P}_{LM} x = 0, \qquad \hat{P}_M y = y$$

and hence, because of the self-adjointness of the projections,

$$(x, y) = (\hat{P}_L x, \hat{P}_M y) = (\hat{P}_M \hat{P}_L x, y) = (\hat{P}_{LM} x, y) = 0$$

namely, $L \obot M$.

Vice-versa, suppose now that $L \obot M$ and take $x \in H$. Since LM is a closed subspace of H, there is one and only one decomposition of x such that

$$x = u + v, \qquad u \in LM, \qquad v \in \overline{LM}$$

Hence $\hat{P}_L u = \hat{P}_M u = u$, since u belongs to both L and M. Moreover, $\hat{P}_L v$ and $\hat{P}_M v$ are in $\overline{LM}$, namely, are orthogonal to LM since, if $w \in LM$, we have

$$(\hat{P}_L v, w) = (v, \hat{P}_L w) = (v, w)$$

(remember that $w \in L$), but $(v, w) = 0$ since $w \in LM$ and $v \in \overline{LM}$; hence $\hat{P}_L v \in \overline{LM}$ (and similarly $\hat{P}_M v \in \overline{LM}$). However, by definition of projection,

[4] See, for example, von Neumann,[1] p. 81.

$\hat{P}_L v \in L$, so that $\hat{P}_L v \in L\overline{LM} = L - M$ (see Remark 1). Since by hypothesis $L - M \perp M$, we have also $\hat{P}_M \hat{P}_L v = \hat{P}_L \hat{P}_M v = 0$ and hence

$$\hat{P}_L \hat{P}_M x = u + \hat{P}_L \hat{P}_M v = u$$

$$\hat{P}_M \hat{P}_L x = u + \hat{P}_M \hat{P}_L v = u$$

namely $[\hat{P}_L, \hat{P}_M] = \hat{0}$. ◇

Remark 4. Formal properties of subspaces are not always immediately transferred into formal properties of projection operators. In fact, if to M corresponds $\hat{P}$, we know[5] that to $\bar{M}$ corresponds $\hat{I} - \hat{P}$ (here $\hat{I} = \hat{P}_H$ is the identity operator). However, the correspondence is not so straightforward in other cases: if to L, M correspond respectively the projections $\hat{P}$, $\hat{Q}$, we cannot expect that to LM simply corresponds $\hat{P}\hat{Q}$, to $L + M$ corresponds $\hat{P} + \hat{Q}$, and to $L - M$ corresponds $\hat{P} - \hat{Q}$. Indeed, we must bear in mind that in general $\hat{P}\hat{Q}$, $\hat{P} + \hat{Q}$, $\hat{P} - \hat{Q}$ are not even projection operators. In fact, we have that $\hat{P}\hat{Q}$ is the projector of LM only when $L \oplus M$, namely (see the previous proposition) when $\hat{P}$ and $\hat{Q}$ commute. Moreover, still supposing that $L \oplus M$, namely that $\hat{P}$ and $\hat{Q}$ commute, from $L + M = \overline{\bar{L}\bar{M}}$ we have immediately that

$$\hat{P}_{L+M} = \hat{P} + \hat{Q} - \hat{P}\hat{Q}$$

so that the projector of $L + M$ is $\hat{P} + \hat{Q}$ only if $\hat{P}\hat{Q} = \hat{0}$ (namely, if they are orthogonal projections). Finally (once more in the same hypothesis of commutative projections) from the results of Remark 1 we have that

$$\hat{P}_{L-M} = \hat{P} - \hat{P}\hat{Q}$$

and coincides with $\hat{P} - \hat{Q}$ only when $\hat{P}\hat{Q} = \hat{Q}$, namely when $M \subseteq L$.

Remark 5. It is easy to see that $L \oplus M$ if and only if $L \oplus \bar{M}$; in fact, from Proposition 2 we know that $L \oplus M$ is equivalent to $[\hat{P}_L, \hat{P}_M] = \hat{0}$ and, since $\hat{P}_{\bar{M}} = \hat{I} - \hat{P}_M$, it is also equivalent to $[\hat{P}_L, \hat{P}_{\bar{M}}] = \hat{0}$; namely (always from Proposition 2), it is equivalent to $L \oplus \bar{M}$. Given the symmetry properties of the relation of w-orthogonality, this means also that $L \oplus M$ is equivalent to $\bar{L} \oplus M$ and $\bar{L} \oplus \bar{M}$.

Remark 6. When $L \oplus M$, we have also that $L - M = L\bar{M}$. Indeed, since $\bar{M} \subseteq \bar{L} + \bar{M} = \overline{LM}$, we have also $L\bar{M} \subseteq L\overline{LM} = L - M$. However, since $L \oplus M$, $L - M = L\overline{LM}$ must be orthogonal to M, namely $L\overline{LM} \subseteq \bar{M}$, and hence $L - M = L\overline{LM} = L(L\overline{LM}) \subseteq L\bar{M}$. As a consequence, $L - M = L\bar{M}$.

[5] See, for example, von Neumann,[1] p. 78.

Remark 7. Two projections $\hat{P}$, $\hat{Q}$ are said to be orthogonal when $\hat{P}\hat{Q}=\hat{0}$. This implies that, in order to give rise to orthogonal projections, two subspaces must be orthogonal (so that $LM=[0]$) and not only w-orthogonal: w-orthogonal subspaces, according to Proposition 2, give rise only to commuting, but not orthogonal, projections.

Definition 3. We say that a family $\mathscr{L}$ of closed subspaces of H is a *lattice of subspaces* when it satisfies the following properties:

(L_1) $H \in \mathscr{L}$, $[0] \in \mathscr{L}$,

(L_2) $L \in \mathscr{L} \Rightarrow \bar{L} \in \mathscr{L}$,

(L_3) $L, M \in \mathscr{L} \Rightarrow LM \in \mathscr{L}$.

Remark 8. The properties L_1, L_2, L_3 imply even that $L+M=\overline{\bar{L}\bar{M}}$ and $L-M=L\overline{LM}$ belong to $\mathscr{L}$, so that it is not necessary to list them among the properties of $\mathscr{L}$.

Remark 9. The properties L_1, L_2, L_3 translate in the language of the subspaces of a Hilbert space the characteristics of the sets constituting an *algebra* in the language of the set theory. In fact, these lattices of subspaces will be supposed to play the same role in representing statements and propositions that the algebras of subsets play in the Kolmogorovian probability.

Example 1. Very simple cases of lattices of subspaces are the following:

$$\mathscr{L}_* = \{H, [0]\}$$
$$\mathscr{L}_L = \{H, L, \bar{L}, [0]\}$$
$$\mathscr{L}^* = \mathscr{S}(H)$$

where $\mathscr{S}(H)$ is the family of all the possible subspaces of H.

Definition 4. We say that a lattice $\mathscr{L}$ is *distributive* when it satisfies the following property:

(L_4) $L, M, N \in \mathscr{L} \Rightarrow L(M+N) = LM + LN$.

Remark 10. The previous definition makes sense since lattices are not in general distributive, as can be seen from a simple counterexample: take in an Euclidean three-dimensional space H the sunspaces M, N consisting of two noncoincident straight lines passing through the origin, and the subspace L consisting in a plane passing through the origin and containing neither M nor N. It is clear now that $M+N$ is a plane cutting L along the one-dimensional line $L(M+N)$, whereas $LM=LN=[0]$ so that $LM+LN=[0]\neq L(M+N)$.

Definition 5. We say that a lattice $\mathscr{L}$ is *orthogonal* when it satisfies the following property:

(L_5) $L, M\in\mathscr{L}\Rightarrow L \circledcirc M$.

Example 2. Let H be a separable Hilbert space and $\{x_n\}_{n\in N}$ a sequence of vectors which is everywhere dense in H (of course, the vectors x_n are not in general orthogonals). Let us now indicate with $M_I=[x_k]_{k\in I}$, $I\subseteq N$ the subspaces spanned by the vectors x_k with the k's belonging to a subset I of N. It is easy to see that the family of subspaces $\{M_I\}_{I\subseteq N}$ is not in general an orthogonal lattice of subspaces for H since (L_2) and (L_4) are not satisfied. However, it can be seen that if $\{x_n\}_{n\in N}$ is a complete orthonormal system, the subspaces M_I become w-orthogonals and $\{M_I\}_{I\subseteq N}$ becomes an orthogonal lattice.

Proposition 3. $\mathscr{L}$ is a distributive lattice if and only if it is an orthogonal lattice.

Proof. Let us suppose firstly that $\mathscr{L}$ is orthogonal: this is equivalent (from Proposition 2) to say that $[\hat{P}_L, \hat{P}_M]=\hat{0}$ for every $L, M\in\mathscr{L}$, and hence that $\hat{P}_L\hat{P}_M=\hat{P}_{LM}$. As a consequence, we have for every $L, M, N\in\mathscr{L}$ that (see also Remark 5)

$$\begin{aligned}\hat{P}_{L(M+N)}&=\hat{P}_L\hat{P}_{M+N}=\hat{P}_L(\hat{P}_M+\hat{P}_N-\hat{P}_M\hat{P}_N)\\&=\hat{P}_L\hat{P}_M+\hat{P}_L\hat{P}_N-\hat{P}_L\hat{P}_M\hat{P}_N=\hat{P}_L\hat{P}_M+\hat{P}_L\hat{P}_N-\hat{P}_L^2\hat{P}_M\hat{P}_N\\&=\hat{P}_{LM}+\hat{P}_{LN}-\hat{P}_{LM}\hat{P}_{LN}=\hat{P}_{LM+LN}\end{aligned}$$

and hence $L(M+N)=LM+LN$, namely, $\mathscr{L}$ is distributive.

Vice-versa, if $\mathscr{L}$ is distributive, for every $L, M\in\mathscr{L}$ we have

$$L-M=L\overline{LM}=L(\bar{L}+\bar{M})=L\bar{M}$$

so that $\mathscr{L}$ is orthogonal since $L\bar{M}\perp M$. □

Remark 11. A consequence of Proposition 3 is that we can arbitrarily use the words *orthogonal* and *distributive* in order to qualify a lattice of subspaces. Hence, in order to simplify our terminology, from now on we will speak only in terms of *orthogonal lattices*.

Proposition 4. For every L, M belonging to an orthogonal lattice $\mathscr{L}$, we have

$$L+M=(L-M)+M$$

Proof. Since $L-M\subseteq L$, we have also that $(L-M)+M\subseteq L+M$, so that it will be enough to show that $L+M\subseteq(L-M)+M$, nemaly, that every $x\in L+M$ can be decomposed in the sum $u+v$ with $u\in M$ and $v\in L-M$. In fact, take $x\in L+M$ and define $u\in M$ as the orthogonal projection of x on M; then define $v=x-u\in\bar{M}$. Of course, $x=u+v$ with $u\in M$, $v\in\bar{M}$. We will show now that, since $\mathscr{L}$ is orthogonal, we have also $v\in L-M$. First of all, we remark that $v=x-u$ belongs to both $\bar{M}$ and $L+M$, namely, that $v\in\bar{M}(L+M)$. However, since $\mathscr{L}$ is orthogonal, we have also, by Proposition 3, that

$$\bar{M}(L+M)=\bar{M}L+\bar{M}M=\bar{M}L$$

Moreover, since $L \oplus M$, we have also $L-M=L\bar{M}$ (see Remark 6), and hence $v\in\bar{L}(L+M)=L\bar{M}=L-M$ as required.

Definition 6. We say that a family $\mathscr{P}$ of projection operators is a *lattice of projections* when it satisfies the following properties:

(P_1) $\hat{P}, \hat{Q}\in\mathscr{P}\Rightarrow[\hat{P},\hat{Q}]=\hat{0}$,

(P_2) $\hat{I}=\hat{P}_H\in\mathscr{P}$, $\hat{0}=\hat{P}_{[0]}\in\mathscr{P}$,

(P_3) $\hat{P}\in\mathscr{P}\Rightarrow\hat{I}-\hat{P}\in\mathscr{P}$,

(P_4) $\hat{P}, \hat{Q}\in\mathscr{P}\Rightarrow\hat{P}\hat{Q}\in\mathscr{P}$.

Remark 12. It is clear, since projections in a lattice $\mathscr{P}$ always commute, that, according to Proposition 2, a lattice of projections $\mathscr{P}$ is in correspondence with one and only one *orthogonal* lattice of subspaces $\mathscr{L}$. On the other hand, there are no lattices of projections corresponding to *nonorthogonal* lattices of sunspaces; in fact, in this case projections no longer commute and hence products of projections are no longer projections, so that we cannot coherently express the property (P_4). This fact is connected with the possibility of having *joint* statements and will be discussed later.

Example 3. Very simple cases of lattices of projections are:

$$\mathscr{P}_*=\{\hat{I},\hat{0}\}$$
$$\mathscr{P}_P=\{\hat{I},\hat{P},\hat{I}-\hat{P},\hat{0}\}$$

Of course, the family of all the projections is not a lattice of projections since they in general do not commute.

Example 4. Given a self-adjoint operator $\hat{X}$, let us consider its *resolution of identity*[6] $\hat{F}_X(x)$ and the projections

$$\hat{F}_X(B) = \int_B d\hat{F}_X(x) = \int_{\mathbf{R}} I_B(x)\, d\hat{F}_X(x) = I_B(\hat{X})$$

where

$$I_B(x) = \begin{cases} 1, & x \in B, \\ 0, & x \in \bar{B}, \end{cases} \qquad B \in \mathscr{B}(\mathbf{R})$$

and $\mathscr{B}(\mathbf{R})$ is the σ-algebra of the Borel sets on the real line $\mathbf{R}$. It is easy to see that the family of projections

$$\mathscr{P}_X = \{\hat{F}_X(B)\}_{B \in \mathscr{B}(\mathbf{R})}$$

is in fact a lattice of projections since its elements always commute and, moreover,

$$\hat{0} = \hat{F}_X(\varnothing),\ \hat{I} = \hat{F}_X(\mathbf{R})$$
$$\hat{I} - \hat{F}_X(B) = \hat{F}_X(\bar{B})$$
$$\hat{F}_X(B_1)\, \hat{F}_X(B_2) = \hat{F}_X(B_1 B_2)$$

belong to $\mathscr{P}_X$ because of the σ-algebra properties of $\mathscr{B}(\mathbf{R})$. In particular, if $B = (a, b]$, we have $\hat{F}_X(a, b] = \hat{F}_X(b) - \hat{F}_X(a)$, so that, if we define

$$F_b(x) = I_{(-\infty, b]}(x) = \begin{cases} 1, & x \leqslant b \\ 0, & x > b \end{cases}$$

we have also $\hat{F}_X(x) = F_x(\hat{X})$. The lattice $\mathscr{P}_X$ will be said *generated* by $\hat{X}$.

Definition 7. We will say that two lattices of projections (orthogonal lattices of subspaces) are *compatible* when every projection of the first commutes with every projection of the second (every subspace of the first is w-orthogonal to every subspace of the second). Moreover, two self-adjoint operators $\hat{X}$, $\hat{Y}$ are said to be compatible when the generated lattices $\mathscr{P}_X$, $\mathscr{P}_Y$ are compatible. In the same sense we can say that an operator $\hat{X}$ is compatible with a lattice $\mathscr{P}$ when the lattices $\mathscr{P}_X$ and $\mathscr{P}$ are compatible.

[6] See, for example, von Neumann,[1] pp. 118 and 252.

Remark 13. When two operators $\hat{X}$, $\hat{Y}$ commute, they are also compatible. Moreover, if $\hat{X}$, $\hat{Y}$ are compatible, we can always find a third operator $\hat{Z}$ and two functions $f(\cdot)$, $g(\cdot)$ such that $\hat{X} = f(\hat{Z})$, $\hat{Y} = g(\hat{Z})$.[7]

Remark 14. It is time to remark now that, despite all the analogies between σ-algebras of events and lattices of projections, there are also fundamental differences which are the basis of the distinction between quantum and classical probability. In a classical world, when we have two different (in the sense that they are not contained one into the other) σ-algebras of events of a set Ω, we can always find a larger σ-algebra containing both. In a sense a σ-algebra represents the set of the statements that we can make in our theory. For example, if $\mathscr{F}_\xi$ is the σ-algebra generated by a random variable ξ, it contains all the statements about ξ of the form $\{\xi \in B\}$, $B \in \mathscr{B}(\mathbf{R})$. If now η is a random variable which is not $\mathscr{F}_\xi$-measurable, this means that in general the statements about η cannot be formulated in $\mathscr{F}_\xi$, namely in terms of ξ. Hence, in a probability space $(\Omega, \mathscr{F}_\xi, \mu)$ we cannot even formulate the question *what is the probability that* η *has a value falling in* $B \in \mathscr{B}(\mathbf{R})$? In fact in this case $\mathscr{F}_\eta$ contains events that do not belong to $\mathscr{F}_\xi$. However, we can always find a larger σ-algebra $\mathscr{F}$ containing both $\mathscr{F}_\xi$ and $\mathscr{F}_\eta$ (for example, the smallest one $\mathscr{F}_{\xi\eta}$) and we can build a new probability space $(\Omega, \mathscr{F}, \mu)$ where all the questions about ξ and η can be formulated. If, on the contrary, η already is $\mathscr{F}_\xi$-measurable, that means that questions about η are in fact questions about ξ: this is clearly expressed in the well known fact that in this case there is a function $f(\cdot)$ such that $\eta = f(\xi)$. These remarks no longer hold in the quantum case: when we have two lattices of projections (or equivalently two orthogonal lattices of subspaces), we cannot always merge them in a larger one because of the commutation rule (P_1) (or equivalently L_5) for subspaces, as can be seen from the following examples.

Example 5. From the following two lattices of projections $\mathscr{P}_P = \{\hat{I}, \hat{0}, \hat{P}, \hat{I} - \hat{P}\}$, $\mathscr{P}_Q = \{\hat{I}, \hat{0}, \hat{Q}, \hat{I} - \hat{Q}\}$, we can build a lattice containing both $\mathscr{P}_P$ and $\mathscr{P}_Q$ if and only if the projection operators $\hat{P}$ and $\hat{Q}$ commute, since the new lattice must contain $\hat{P}$ and $\hat{Q}$, and (P_1) must hold: namely, the two lattices must be compatible.

Example 6. The quantum analog of the classical case of two σ-algebras $\mathscr{F}_\xi$ and $\mathscr{F}_\eta$ is the case of two lattices of projections $\mathscr{P}_X = \{\hat{F}_X(A)\}_{A \in \mathscr{B}(R)}$ and $\mathscr{P}_Y = \{\hat{F}_Y(B)\}_{B \in \mathscr{B}(R)}$ generated by two slf-adjoint operators $\hat{X}$, $\hat{Y}$: from remarks similar to that of Example 5, we conclude

[7] For more details, see von Neumann,[1] pp. 172-173.

immediately that a larger lattice containing both $\mathscr{P}_X$, $\mathscr{P}_Y$ exists only if $\hat{X}$ and $\hat{Y}$ are compatible (for example, if $[\hat{X}, \hat{Y}] = \hat{0}$).

Remark 14 (continuation). As a consequence, it can be said that a crucial difference between classical and quantum probability models lies in the fact that, while in the classical case we can speak in terms of set theory, in the quantum case we speak in terms of Hilbert spaces, namely of vector spaces. This means that, in order to represent statements and propositions, the σ-algebras of subsets (σ-algebras of indicators) must be replaced by orthogonal lattices of subspaces (lattices of projections) and this requirement introduces in the theory a fundamental rigidity [expressed in the properties (L_5) or (P_1)] with far reaching consequences on the possibility of making joint statements, as we will see later. It is clear now that a central role in this situation is played by the *superposition principle* which can be considered as the main reason for working in vector spaces.[8]

Definition 7. A *quantum space* (QS) is a couple $(H, \mathscr{S}(H))$, where H is a Hilbert space and $\mathscr{S}(H)$ is the (nonorthogonal) lattice of all its subspaces. A *quantum representation* (QR) is a couple $(H, \mathscr{L})$, where $\mathscr{L}$ is an orthogonal lattice of subspaces of the Hilbert space H.

Definition 8. A *quantum probability measure* (QPM) defined on a (not necessarily orthogonal) lattice $\mathscr{L}$ is an application λ: $\mathscr{L} \to [0, 1]$ such that $\lambda(H) = 1$ and that λ is *additive*, namely, that for every L, M belonging to $\mathscr{L}$ we have

$$L \perp M \Rightarrow \lambda(L + M) = \lambda(L) + \lambda(M)$$

Definition 9. A *quantum probability space* (QPS) is a triple $(H, \mathscr{S}(H), \lambda)$, where H is a Hilbert space, $\mathscr{S}(H)$ the lattice of all its subspaces, and λ a QPM defined on $\mathscr{S}(H)$. A *quantum probability representation* (QPR) is a triple $(H, \mathscr{L}, \lambda)$, where H is a Hilbert space, $\mathscr{L}$ an orthogonal lattice of subspaces of H, and λ a QPM defined on $\mathscr{L}$.

Remark 15. In a QPS, orthogonal subspaces play a role similar to that of disjoint subsets in classical probability spaces. Moreover, it is easy to verify that other properties are similar to that of the classical case. For example, from the additivity of λ we can deduce that $\lambda(\bar{L}) = 1 - \lambda(L)$ and that $\lambda([0]) = 0$. The folowing simple propositions constitute further examples.

[8] See, for example, P. A. M. Dirac,[2] Chap. 1.

Proposition 5. In a QPS, for every $L, M \in \mathscr{S}(H)$ we have

$$M \subseteq L \Rightarrow \lambda(L-M) = \lambda(L) - \lambda(M)$$

Proof. Since $M \subseteq L$, we have also that $L - M = L\bar{M} \perp M$ (see also Remark 1). Moreover, it is immediately seen that $(L-M)+M = L$, so that from the additivity we get

$$\lambda(L) = \lambda[(L-M)+M] = \lambda(L-M) + \lambda(M)$$

and hence $\lambda(L-M) = \lambda(L) - \lambda(M)$. □

Corollary 1. In a QPS, for every $L, M \in \mathscr{S}(H)$ we have

$$M \subseteq L \Rightarrow \lambda(M) \leqslant \lambda(L)$$

Proof. It follows from Proposition 5 if we take into account the positivity of λ. □

Corollary 2. In a QPS, for every $L, M \in \mathscr{S}(H)$ we have

$$\lambda(L-M) = \lambda(L) - \lambda(LM)$$

Proof. Since $LM \subseteq L$, we have from Proposition 5 that

$$\lambda(L-M) = \lambda(L\overline{LM}) = \lambda[L\overline{L(LM)}] = \lambda(L-LM) = \lambda(L) - \lambda(LM) \quad \square$$

Proposition 6. In a QPS, for every $L, M \in \mathscr{S}(H)$ we have

$$L \oplus M \Rightarrow \lambda(L+M) = \lambda(L) + \lambda(M) - \lambda(LM)$$

Proof. Since $L \oplus M$, it follows from Proposition 4 that $L+M = (L-M)+M$, and from Definition 2 that $L-M \perp M$. Hence, from the additivity of λ, we have that

$$\lambda(L+M) = \lambda[(L-M)+M] = \lambda(L-M) + \lambda(M)$$

and the proposition follows from Corollary 2. □

Remark 16. The property of Proposition 6 is very similar to the analogous property of the classical case. However, like many other properties, it does not hold universally in a QPS since, in general, two subspaces are not w-orthogonal. Only in a particular QPR is this property always true, since in a QPR all the subspaces are w-orthogonal.

Remark 17. A similar situation arises for the so-called *total probability theorem* of the classical case. In fact, it is very easy to see that an analogous property holds in the quantum case only in every particular QPR, but not in an entire QPS since the lattice $\mathscr{S}(H)$ is not distributive. To see it, let us consider a *finite decomposition* of H, namely, a finite family of subspaces $\{M_k\}_{k=1,\dots,n}$ such that

$$M_j \perp M_k, \qquad j \neq k, \qquad \sum_{k=1}^{n} M_k = H$$

If now L is a subspace of H, since $\mathscr{S}(H)$ is not a distributive lattice we have that

$$L = L \sum_{k=1}^{n} M_k \neq \sum_{k=1}^{n} LM_k$$

and hence

$$\lambda(L) = \lambda\left(L \sum_{k=1}^{n} M_k\right) \neq \lambda\left(\sum_{k=1}^{n} LM_k\right) = \sum_{k=1}^{n} \lambda(LM_k)$$

namely, the total probability theorem does not hold. In particular we have

$$\lambda(L) \neq \lambda(LM) + \lambda(L\bar{M})$$

Of course, if L and M_k's are all elements of the same orthogonal lattice (for example, if we restrict ourselves to a particular QPR), the intersection of subspaces is distributive and the total probability formula is valid.

Remark 18. The notion of QPS and QPR can also in a very natural way be discussed in terms of projections since they are in a one-to-one correspondence with closed subspaces of H. However, we must remark that $\mathscr{P}(H)$, defined as the family of all the projections corresponding to the closed subspaces belonging to $\mathscr{S}(H)$, is not a lattice of projections since $\mathscr{S}(H)$ is not an orthogonal lattice of subspaces (see Remark 12 and Example 3). That notwithstanding, it is possible to give a definition of QPS (and of QPR) based on projections through the notion of *statistical operator*[9] $\hat{U}$, since it is easy to see that the application $\nu: \mathscr{S}(H) \to [0, 1]$ defined as

$$\nu(L) = \mathrm{Tr}(\hat{U}\hat{P}_L), \qquad L \in \mathscr{S}(H)$$

[9] For more details see, for example, von Neumann,[(1)] Chap. IV.

[10] See, for example, von Neumann,[(1)] p. 81.

is a QPM. In fact, to verify the additivity it is enough to remark that[10] when $L \perp M$ we have $\hat{P}_{L+M} = \hat{P}_L + \hat{P}_M$. Hence, we can give the following alternative definitions of QPS and QPR:

Definition 10. A QPS is a triple $(H, \mathscr{P}(H), \hat{U})$, where $\mathscr{P}(H)$ is the family of all the projections defined on a Hilbert space H and $\hat{U}$ is a statistical operator. A QPR is a triple $(H, \mathscr{P}, \hat{U})$, where $\mathscr{P}$ is a lattice of projections defined on H and $\hat{U}$ is a statistical operator.

Remark 19. Given a QPS (QPR) in the sense of Definition 10, it is clear that a unique corresponding QPS (QPR) in the sense of Definition 9 always exists. To establish the inverse statement, however, is a much more complicated task. This will amount to an answer to the following question: Given a QPS $(H, \mathscr{S}(H), \lambda)$ (in the sense of Definition 9), is it always possible to uniquely determine a statistical operator $\hat{U}$ such that

$$\lambda(L) = \mathrm{Tr}(\hat{U}\hat{P}_L), \qquad L \in \mathscr{S}(H)$$

so that the QPS $(H, \mathscr{P}(H), \hat{U})$ (in the sense of Definition 10) can be rightly considered as corresponding to $(H, \mathscr{S}(H), \lambda)$? Gleason has shown[(3)] that the answer to this question is positive if the dimension of H is at least 3. Counterexamples can be given if the dimension of H is less than 3 (see Appendix).

Remark 20. Coming now to the problem of defining *joint statements*, it is clear that this is not allowed in general in a QPS $(H, \mathscr{P}(H), \hat{U})$, even if it becomes possible in every particular QPR $(H, \mathscr{P}, \hat{U})$ (see Remark 12). In fact, if $\hat{P}, \hat{Q}$ belong to $\mathscr{P}(H)$, in general they do not commute, so that $\hat{P}\hat{Q}$ is not a new projection (it is not even self-adjoint) and no other combination can be found to play the role of the representative of joint statements. For example, the operator

$$\frac{1}{2}\{\hat{P}, \hat{Q}\} = \frac{\hat{P}\hat{Q} + \hat{Q}\hat{P}}{2}$$

is self-adjoint but still it is not a projection. Finally, we could hope to define joint statements working in a QPS in the sense of Definition 9: Given $\hat{P}, \hat{Q} \in \mathscr{P}(H)$, if $L, M \in \mathscr{S}(H)$ are the corresponding subspaces, we could think of defining the joint statement as $\hat{P}_{LM}$, which is now a projection operator. However, as already remarked in a previous paper,[(4)] the fact that the total probability theorem does not hold (see Remark 17) prevents us from coherently using the intersections LM as representatives of joint statements in a general QPS. For example, $\lambda(L) \neq \lambda(LM) + \lambda(L\bar{M})$, and hence we cannot consider LM and $L\bar{M}$ as true joint statements since they should be *alternative* statements and hence $\lambda(L) = \lambda(LM) + \lambda(L\bar{M})$ should logically be verified. As a consequence of this impossibility of a definition of joint statements, we must also remark here that in a QPS it is impossible to define *independent statements* and hence *independent observables* in the general case: In a QPS only the notion of *noncorrelation* can be coherently given, but not that of independence.

Remark 21. The previous remarks about the impossibility of coherently defining joint statements in a QPS forbids also a definition of *conditional probabilities* along the usual lines (see also Ref. 4 for further details). For example, given a subspace M such that $\lambda(M) \neq 0$, we could try to define a conditional probability as

$$\lambda(L \mid M) = \frac{\lambda(LM)}{\lambda(M)} = \frac{\mathrm{Tr}(\hat{U}\hat{P}_{LM})}{\mathrm{Tr}(\hat{U}\hat{P}_{M})}$$

However, $\lambda(\cdot \mid M)$ should behave like a new QPM, namely, it should be additive, but it is easy to see that $\lambda(\cdot \mid M)$ is not additive since $\mathscr{S}(H)$ is not distributive:

$$\lambda(L \mid M) + \lambda(\bar{L} \mid M) = \frac{\lambda(LM) + \lambda(\bar{L}M)}{\lambda(M)} = \frac{\lambda(LM + \bar{L}M)}{\lambda(M)}$$

$$\neq \frac{\lambda[(L + \bar{L})\, M]}{\lambda(M)} = 1$$

Hence we cannot consider $\lambda(\cdot \mid M)$ a conditional probability since it is not a QPM. As a consequence, if we want to define conditional probabilities, we must follow a (well-known in quantum mechanics) different road which has nothing to do with joint statements; rather, it is connected with the idea of a sequence of preparations. In this way it can be seen that cannot always retain, after the second measurement, the information obtained from the first preparation—another behavior of the quantum physics in striking contrast with the classical case (for a more detailed discussion see, for example, Ref. 4).

3. CONCLUSIONS

A striking difference between quantum and classical probability becomes apparent in the structure of the mathematical models that can be built in order to describe the sets of statistical data in the two situations. In every such mathematical model a central aim is to reduce to a sort of unity the multiplicity of the world of the measurements. Here, by the word *unity* we mean the individuation of a sort of common source of all the probabilistic statements which are possible in a theory (after all, the *calculus of the probabilities* has always been connected with the idea of calculating the probabilities in complicated situations from that of simple events). The difference between the two approaches is in the way in which different informations (coming from different measurements) are brought together... if they can. In the classical world this problem has an answer in the construction of the usual Kolmogorovian models $(\Omega, \mathscr{F}, \mu)$; however, we should realize that, even in the classical case, this model is not given *a priori* but is, in a sense, the result of an operation of *unification and abstraction* executed on the set of our empirical data: the physical quantities X, Y,... can be endowed with individual distributions; then we consider their joint (or conditional) distributions; and finally we can give an abstract formulation by defining the space of all the possible results of individual or joint measurements. This leads to the idea of the *sample space* Ω, whose elements are exactly the abstraction of the possible results of an experiment. As a consequence of this operation, all the richness of the empirical world is unified in the mathematical model $(\Omega, \mathscr{F}, \mu)$, in the sense that every probabilistic statement (probabilities of events, distribution functions, joint and conditional probabilities...) concerning the physical quantities X, Y,... can now be deduced from a unique probability measure μ by means of the idea of the random variables as measurable functions ξ, η,... defined on Ω with values in some other measurable space. Moreover, the model is in some sense indefinitely open. Take a new measurable quantity Z (not initially considered for our construction even if we suppose, for the sake of simplicity, that Ω is rich enough to contain also the results of the measurements of Z): if it can be described by a random variable ζ which is already measurable with respect to the σ-algebra $\mathscr{F}$, nothing more is needed. If, on the contrary, this is not so (namely if $\mathscr{F}_\zeta$ is not contained in $\mathscr{F}$), it is always possible to widen the σ-algebra $\mathscr{F}$ in a σ-algebra containing both $\mathscr{F}_\zeta$ and $\mathscr{F}$. This means that one can always describe all the statistical data in this unified way.

So strong is the suggestion of this picture that we are often led to think that this model is in fact coincident with the underlying reality of the physical world, that this reality can be asserted only if such a $(\Omega, \mathscr{F}, \mu)$ can be constructed and we forget that, on the contrary, this model could be nothing more than a useful abstraction. This can be considered the reason for the fact that sometimes we consider paradoxical a situation in which such a picture is not possible, even if other unified models are available, and, as already seen in the Remark 14, this is exactly the situation in the quantum physics.[5]

In fact, the *unification plus abstraction* operation is possible also in the quantum world, but the result is not a Kolmogorovian model; rather, it is a QPS $(H, \mathscr{S}(H), \mu)$ (or $(H, \mathscr{P}(H), \hat{U})$). In some sense this mathematical model does exactly the same job as the Kolmogorovian models do in the classical world: the probabilities of all the possible statistical statements (however, not all the *joint* statements are possible: see Remark 20) can be derived from this unified model. But this model cannot be put in correspondence with a unique classical model whatsoever. To see that, let us remark that, given a QPR $(H, \mathscr{P}, \hat{U})$ (for example, let $\mathscr{P}$ be the lattice of projections generated by a complete set of commuting observables[11]), there is always a way to describe the same empirical results in a perfectly classical model. However (see Remark 14), as soon as we introduce in our description a new observable noncompatible with $\mathscr{P}$, we are obliged to abandon our QPR and work in a QPS $(H, \mathscr{P}(H), \hat{U})$, where $\mathscr{P}(H)$ is no longer a lattice of projection, since no lattice can contain noncommuting projections. In this case we immediately lose the possibility of describing our empirical data in a classical Kolmogorovian model. In a sense, we could say that a quantum probabilistic model is something allowing of many partial (namely, describing only some of the physical observables) but of no unified classical model, the different partial classical models being not mutually compatible.

Classical models can actually be constructed in every QPR for our quantum system. For example, it is very well known that for a quantum particle, in the representation of the position, we can build a completely classical theory (the stochastic mechanics[6]) which perfectly simulates all the quantum results. In this theory there are classical behaviors which are not compatible with the quantum mechanics, but they are not directly observable in the sense that they correspond to *virtual* (namely, not satisfying the stochastic variational principle which generates the dynamics of the theory) stochastic processes. Even a transformation theory is allowed in the stochastic mechanics[7] so that we can also change our representation and consequently locally *explore* our QPS by meansof classical models which are observably not distinguishable from it.

We must finally remark that in a QPS, namely, in a coherent unified mathematical model for a set of quantum statistical data, the notion of *conditioning* as acquisition of new information is radically different from the corresponding notion in the classical world. As already pointed out (see Ref. 4 and refernces therein), while in the classical physics this acquisition

[11] See, for example, P. A. M. Dirac,[2] Chap. 3, §14.

process is always cumulative, in quantum physics it is not since there are incompatible informations (connected to the different QPR's of a unique QPS), which can be considered as a probabilistic way to state the uncertainty principle. This remark, along with that about the superposition principle (see Remark 14), puts the formal structure of a QPS in a direct connection with the basic postulates of the quantum mechanics. But there is more: the new way of calculating the conditional probabilities in a QPS (namely, the calculation of transition probabilities such as $|(x, y)|^2$ for $x, y \in H$) puts in evidence a limitation of the classical definition of conditional probability which was already remarked, although in a different context, by De Finetti:[12] in the classical way of calculating conditional probabilities, the new information does not in fact change our previous evaluations of the likelihood of events: the definition

$$\mu(A \mid B) = \frac{\mu(A \cap B)}{\mu(B)}$$

is inherently based on the *a priori* probability $\mu(\cdot)$ and "derives indeed from the same *a priori* judgment, by subtracting, let us say, the doubt components concerning the trials whose result has been acquired."(8) QPS's belong, instead, to a different class of mathematical models: those where a conditioning can modify even the *a priori* probability measure.

ACKNOWLEDGMENTS

The author wants to thank Professors L. Accardi, F. Guerra, and J. P. Vigier for invaluable discussions and suggestions.

APPENDIX

We will continue here the discussion started at the end of Remark 19 in order to show that it is not always possible to realize a QPM through a statistical operator $\hat{U}$. Let us consider indeed a complex, two-dimensional Hilbert space H with the usual norm $\|x\|^2 = (x, x)$. The unique subspace of dimension 2 is of course H itself; the unique subspace of dimension 0 is $[0]$. On the other hand, if L is a one-dimensional subspace of H, for every $x, y \in L$ we will have $y = cx$, $c \in \mathbf{C}$: in fact, to be in the same one-dimensional subspace amounts, for two vectors of H, to the equivalence relation of being parallel. In order to choose a unique representative x for

[12] See, for example, Ref. 8, p. 25.

every one-dimensional subspace L (and we will write $L=[x]$), we take a (complete) system of two orthonormal vectors x_1, x_2 ($\|x_1\|^2=\|x_2\|^2=1$, $(x_1, x_2)=0$), and then, if $x_2 \in L$ we take x_2 as the representative of L ($L=[x_2]$). If, on the other hand, $x_2 \notin L$, we consider an arbitrary $y \in L$ and then we will take as representative the vector

$$x=\frac{e^{-i\arg(x_1, y)}}{\|y\|}y, \qquad \arg(x_1, y) \in [0, 2\pi)$$

which is normalized and is characterized by the fact that (x_1, x) is real and $(x_1, x) \in (0, 1]$ (remember that in general $|(x_1, x)| \leqslant \|x_1\| \cdot \|x\| = 1$). In particular, when $x_1 \in L$, then $y=cx_1$, so that the representative chosen will be exactly x_1. This procedure to choose a representative leads, of course, to a unique result. This is obvious if $x_2 \in L$ since, by definition, the unique choice is x_2. On the other hand, if $x_2 \notin L$, suppose that x is the representative of L, namely $L=[x]$ with $\|x\|=1$, $0<(x, x_1)=(x_1, x) \leqslant 1$; we will show that every $y \in L$ with the same characteristics ($\|y\|=1$, $0<(y, x_1)=(x_1, y) \leqslant 1$) must coincide with x. In fact, since $y \in L$ we have $y=cx$; but we must have $|c|=1$ (from $\|y\|=1$) and $\arg(x_1, y)=0$ (in order to keep (x_1, y) real and positive). Hence, $y=x$.

We have shown that every one-dimensional subspace of H can be characterized by means of one and only one normalized vector: we will try now to use this fact in order to describe the set of the one-dimensional subspaces by means of numerical parameters. Let us remember that every $x \in H$ has a unique representation of the form

$$x=c_1x_1+c_2x_2, \qquad c_1=(x_1, x), c_2=(x_2, x)$$

Hence the representatives of our one-dimensional subspaces can be parametrized by means of the complex numbers c_1, c_2; or, better still, since the representatives of subspaces must be normalized and must have c_1 real and positive, our subspaces can be parametrized by means of c_2 only. In fact, when $L=[x_2]$, we will have $c_2=1$ (and $c_1=0$). If, on the contrary, $x_2 \notin L$, we calculate $c_2=(x_2, x)$, and the real and positive number c_1 will be fixed as $c_1=\sqrt{1-|c_2|^2}$. Of course, in this second case we will always have $|c_2|<1$ (and hence $0<c_1 \leqslant 1$). Hence, if we define the following set of complex numbers:

$$\Lambda=\{1\} \cup \{c \in \mathbf{C}: |c|<1\}$$

we have a one-to-one correspondence between the one-dimensional subspaces of H and the elements c of Λ. If we indicate with L_c the subspace individuated by $c \in \Lambda$, we will have, for example, $[x_2]=L_1$, $[x_1]=L_0$, and in general $L_c=[\sqrt{1-|c|^2}\, x_1+cx_2]$.

With this notation the lattice of all the subspaces of H is

$$\mathscr{S}(H) = \{[0], \{L_c\}_{c \in \Lambda}, H\}$$

On the other hand, every arbitrary orthogonal lattice of subspaces of H will be selected by a vector $x \in H$ and will contain the following elements:

$$\mathscr{L}_x = \{[0], [x], [\bar{x}], H\}$$

For example, $\mathscr{L}_{x_1} = \mathscr{L}_{x_2} = \{[0], [x_1], [x_2], H\}$ is a particular orthogonal lattice of subspaces of H.

If now $\mathscr{D} = \{L_c\}_{c \in \Lambda}$ is the family of all the one-dimensional subspaces of H, it is easy to see that we can separate it in two subfamilies $\mathscr{D}_1$, $\mathscr{D}_2$ with the following properties:

(a) every one-dimensional subspace of H is either in $\mathscr{D}_1$ or in $\mathscr{D}_2$, but never in both;

(b) two subspaces of $\mathscr{D}_1$ ($\mathscr{D}_2$) are never mutually orthogonal;

(c) for every subspace of $\mathscr{D}_1$ ($\mathscr{D}_2$) the corresponding orthogonal subspace is in $\mathscr{D}_2$ ($\mathscr{D}_1$).

In fact, two one-dimensional subspaces are orthogonal, $[x] \perp [y]$, if and only if the corresponding complex parameters in Λ, $c = |c|\, e^{i\gamma}$ for x and $b = |b|\, e^{i\beta}$ for y, satisfy the following relations:

$$|b| = \sqrt{1 - |c|^2}, \qquad \beta = \begin{cases} \gamma + \pi, & \text{if} \quad \gamma \in [0, \pi) \\ \gamma - \pi, & \text{if} \quad \gamma \in [\pi, 2\pi) \end{cases}$$

As a consequence, if we define the following subsets of Λ:

$$\Lambda_1 = \left\{c \in \mathbf{C}\colon |c| < \frac{1}{\sqrt{2}}\right\} \cup \left\{c \in \mathbf{C}\colon |c| = \frac{1}{\sqrt{2}};\ \arg c \in [0, \pi)\right\}$$

$$\Lambda_2 = \{1\} \cup \left\{c \in \mathbf{C}\colon \frac{1}{\sqrt{2}} < |c| < 1\right\} \cup \left\{c \in \mathbf{C}\colon |c| = \frac{1}{\sqrt{2}},\ \arg c \in [\pi, 2\pi)\right\}$$

we immediately see that $\Lambda_1 \cup \Lambda_2 = \Lambda$, $\Lambda_1 \cap \Lambda_2 = \varnothing$, that two one-dimensional subspaces represented in Λ_1 (Λ_2) are never mutually orthogonals, and that for every subspace represented in Λ_1 (Λ_2) the corresponding orthogonal vector lies in Λ_2 (Λ_1). Hence, we obtain the required partition of $\mathscr{D}$ in $\mathscr{D}_1$ and $\mathscr{D}_2$ if we define

$$\mathscr{D}_1 = \{L_c\}_{c \in \Lambda_1}, \qquad \mathscr{D}_2 = \{L_c\}_{c \in \Lambda_2}$$

It is immediate to see that $x_1 \in \mathscr{D}_1$ and $x_2 \in \mathscr{D}_2$.

We can now define a simple QPS $(H, \mathscr{S}(H), \lambda)$ by means of the following QPM:

$$\lambda(L) = \begin{cases} 0, & \text{if } L = [0] \\ 1, & \text{if } L = H \\ p, & \text{if } L \in \mathscr{D}_1 \\ 1-p, & \text{if } L \in \mathscr{D}_2 \end{cases}$$

Hence, for example, $\lambda([x_1]) = \lambda(L_0) = p$ and $\lambda([x_2]) = \lambda(L_1) = 1-p$. It is very easy to verify that this λ is additive since if $L \in \mathscr{D}_1(\mathscr{D}_2)$ the unique orthogonal one-dimensional subspace M will belong to $\mathscr{D}_2(\mathscr{D}_1)$ and hence their respective probabilities are p and $1-p$ (or vice-versa), so that $\lambda(L+M) = \lambda(L) + \lambda(M)$.

We will show now that it is impossible to find a statistical operator $\hat{U}$ such that

$$\lambda(L) = \mathrm{Tr}(\hat{U}\hat{P}_L), \qquad L \in \mathscr{S}(H)$$

In fact, in the representation established by means of the orthonormal basis x_1, x_2, we should calculate the four complex components of the 2×2 matrix

$$U_{ij} = (x_i, \hat{U}x_j), \qquad i, j = 1, 2$$

Since for every one-dimensional subspace $L = [x] \in \mathscr{D}$ we have $\mathrm{Tr}(\hat{U}\hat{P}_L) = (x, \hat{U}x)$, we get immediately that

$$U_{11} = p, \qquad U_{22} = 1 - p$$

so that the requirement Tr $\hat{U} = 1$ is authomatically satisfied. Moreover, since $\hat{U}$ must be Hermitean, we have $U_{12} = U_{21}^* = |u|\, e^{i\gamma}$. Finally, $\hat{U}$ must be positive definite and hence, given an arbitrary normalized vector $x = ax_1 + bx_2 \in H$, with $a = |a|\, e^{i\alpha}$, $b = |b|\, e^{i\beta}$, $|a|^2 + |b|^2 = 1$, we must require that

$$(x, \hat{U}x) = (1 - |b|^2)\, p + |b|^2\, (1-p) + 2\, |b| \cdot |u|\, \sqrt{1 - |b|^2}\, \cos(\gamma - \alpha - \beta) \geqslant 0$$

This means that we should have

$$|u| \leqslant \frac{p + (1-2p)\, |b|^2}{2\, |b|\, \sqrt{1 - |b|^2}}$$

for an arbitrary $|b| \in [0, 1]$, and this is possible only if $|u|$ is smaller than the minimum value taken by the right-hand side. It is easy to see that this implies the following limitation:

$$|u| \leqslant \sqrt{p(1-p)}$$

However, we can show that this requirement is incompatible with the values that must be imposed in order to reproduce λ. In fact, take, for example, the two normalized vectors

$$y = \frac{x_1 + x_2}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \qquad z = \frac{x_1 + ix_2}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

Since these vectors correspond respectively to the values $1/\sqrt{2}$ and $i/\sqrt{2}$ of the parameter c, and since these values both lie in Δ_1, by definition we have $\lambda([y]) = \lambda([z]) = p$. If we impose now that

$$(y, \hat{U}y) = (z, \hat{U}z) = p$$

we deduce that $U_{12} = U_{21}^* = \frac{1}{2}(2p-1)(1-i)$ so that $|u| = (1/\sqrt{2})\,|2p-1|$, which is not compatible with the limitation imposed by the positivity requirement for every value of $p \in [0, 1]$. For example, if $p = 1/6$, we get $|u| = \sqrt{8}/6 \geqslant \sqrt{5}/6 = \sqrt{p(1-p)}$. Hence, it is not always possible to reproduce a QPM defined on a two-dimensional QPS by means of a suitable statistical operator $\hat{U}$.

REFERENCES

1. J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).
2. P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1958).
3. A. M. Gleason, *J. Math. Mech.* **6** (1957) 885; E. G. Beltrametti and G. Cassinelli, *Riv. Nuovo Cimento*, **6**, 321 (1976).
4. N. Cufaro-Petroni, *Phys. Lett. A* **160**, 107 (1991); N. Cufaro-Petroni and J. P. Vigier, *Found. Phys.* **22**, 1 (1992).
5. L. Accardi and A. Fedullo, *Lett. N. Cimento* **34**, 161 (1982); L. Accardi, *Phys. Rep.* **77**, 169 (1981).
6. E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton University Press, Princeton, 1967); *Quantum Fluctuations* (Princeton University Press, Princeton, 1985); F. Guerra, *Phys. Rep.* **77**, 263 (1981).
7. F. Guerra and R. Marra, *Phys. Rev. D* **29**, 1647 (1984).
8. B. De Finetti, *Ann. Inst. H. Poincaré* **VII**, 1 (1937).

Printed in Belgium
Verantwoordelijke uitgever:
Hubert Van Maele
Altenastraat 20 - B-8310 St.-Kruis