Foundations of non-commutative probability theory (Extended abstract) *

Daniel Lehmann
School of Engineering and
Center for the Study of Rationality
Hebrew University, Jerusalem 91904, Israel

Abstract

Kolmogorov's setting for probability theory is given an original generalization to account for probabilities arising from Quantum Mechanics. The sample space has a central role in this presentation and random variables, i.e., observables, are defined in a natural way. The mystery presented by the algebraic equations satisfied by (non-commuting) observables that cannot be observed in the same states is elucidated.

1 Introduction

In Quantum Physics a state of a physical system defines random variables corresponding to *observables* that are represented by Hermitian operators. These random variables cannot be treated in the framework, laid down by Kolmogorov in the 30's, which is now standard in probability theory. The main reason is that, in the standard treatment, real random variables are functions from the sample space to the set of reals, implying that all points of the sample space *possess* values for any random variable, whereas the standard understanding of Quantum Physics requires that random variables that correspond to non-commuting operators cannot both have a value at the same time.

This paper proposes a generalization of Kolmogorov's framework that encompasses the non-commuting probabilities arising from Quantum Physics. Contrary to previous efforts, known under the general term of Quantum Logic and which [2] surveys in an authoritative way, in which the sample space is absent, this effort gives a central role to the sample space.

*This work was partially supported by the Jean and Helene Alfassa fund for research in Artifi cial Intelligence

TARK '09, July 6-8, 2009, California
ISBN: 978-1-60558-560-4...$10.00

2 Kolmogorov's setting

We shall recall the now classical setting laid down by Kolmogorov. The description below is not the most succinct possible, but the reader will have no problem showing it is equivalent to his/her favorite presentation.

We start with an arbitrary non-empty set Ω, the *sample space*, whose elements are called *points*.

Definition 1 *A set $\mathcal{F} \subseteq 2^{\Omega}$ of sets of sample points is a σ-field iff it satisfies*

1. $\emptyset \in \mathcal{F}$,
2. *for any $A \in \mathcal{F}$, the complement of A, $A^c = \Omega - A$ is a member of $\mathcal{F}$,*
3. *for any finite or* countably infinite *sequence A_i, $i \in I$ of* pairwise disjoint *elements of $\mathcal{F}$ (for any $i, j \neq i$, $A_i \cap A_j = \emptyset$) their union $\bigcup_{i \in I} A_i$ is a member of $\mathcal{F}$.*

The elements of $\mathcal{F}$ are called events.

Definition 2 *A probability measure (or, distribution) p is a function $p : \mathcal{F} \longrightarrow [0, +\infty]$ such that:*

1. $p(\emptyset) = 0$,
2. $p(\Omega) = 1$,
3. *for any finite or* countably infinite *sequence A_i, $i \in I$ of* pairwise disjoint *elements of $\mathcal{F}$, we have $p(\bigcup_{i \in I} A_i) = \sum_{i \in I} p(A_i)$.*

The definition of a random variable is now the following.

Definition 3 *A random variable of a σ-field $\langle \Omega, \mathcal{F} \rangle$ to a σ-field $\langle \mathcal{R}, \Sigma \rangle$ is a* measurable *function $X : \Omega \longrightarrow \mathcal{R}$, that is to say a function X such that the inverse image by X of any event of Σ is an event of $\mathcal{F}$.*

3 Extant work

The foundation of Quantum Logic was laid by Birkhoff and von Neumann in [1] which set the frame for later work in Quantum Logic. This frame is based on the classical views that quantic propositions are either true or false, that propositions can be composed using negation, conjunction and disjunction, and that the structure to be studied is the consequence relation: which propositions follow from other propositions or sets of propositions. The algebraic structure of such propositions is naturally seen to be an ordered structure, in fact a lattice. Birkhoff and von Neumann noticed that the lattice of interest is not, in general, distributive. Quantum Logic therefore studied non-distributive complemented lattices, satisfying a property weaker than distributivity: modularity was advocated by [1] but most researchers opted for the even weaker orthomodularity.

The probabilistic aspect of Quantum Physics is probably its most revolutionary feature. There is no doubt that a physicist will consider the fact that, in Quantum Physics, a state can define, even in principle, only the probability of observations as more immediately revolutionary than the fact that disjunction does not distribute over conjunction. We shall now describe the way Quantum Logic deals with probabilities. Its analysis of classical probabilities relies on the observation that a σ-field defines a Boolean algebra with countable l.u.b's. A (classical) probability measure is therefore a function that attaches a real number (its probability) to every element of a Boolean algebra and satisfies certain conditions. The concrete algebra of subsets presented in Kolmogorov's setting is replaced by an abstract Boolean algebra. By Stone's representation theorem, there is no loss here since any Boolean algebra is isomorphic to a concrete algebra of subsets. Probability measures in Quantum Logic are therefore analyzed as functions assigning a probability to every element of an orthomodular lattice that satisfy certain properties. But orthomodular (or modular) lattices are not, in general, lattices of sets and the sample space disappears from the picture. This has three serious drawbacks. First the intuitive idea that probability of an event is, in some sense, the measure of the "size" of a set of possibilities cannot be carried on. Secondly, the definition of a random variable, which requires a sample space, is not possible. Thirdly, the special case of classical probabilities is characterized by the boolean character of the lattice and this may seem at odds with the view generally held by physicists that classical physics is the special case of quantum physics in which all operators commute: it is difficult to see boolean lattices as *commutative orthomodular lattices*. A family of algebras generalizing boolean algebras has been proposed in [3] and boolean algebras are exactly the commutative algebras of the family. The relation of those algebras to the present work needs further study.

The first concern has been addressed by setting additional requirements, concerned with Atomicity and Covering, on the lattice structure: see for example axioms H1 and H2 in [8]. Such properties are *not* satisfied in Boolean algebras and therefore classical probabilities are *not* a special case of Quantum probabilities. Random variables may then be defined by functions on the atoms of the structure.

This work proposes a framework for probability theory that generalizes Kolmogorov's and that encompasses Quantum Probability. Classical probability is a special case of Quantum Probability. The sample space is not eliminated, but given some additional structure: it is an Similarity-Projection (SP) structure. These have been defined and studied in [5]. They abstract from the real scalar product in Hilbert spaces.

4 A more general setting

We shall generalize Kolmogorov's setting by assuming some structure on the sample space Ω. We assume there is a *similarity* function $s : \Omega \times \Omega \longrightarrow \mathcal{R}$ that associates a real number, their similarity, to any two sample points. Think of x and y as unitary vectors in a Hilbert space and think of $s(x, y)$ as their real scalar product squared: $\mid \langle x, y \rangle \mid^2$. We shall assume that the pair $\langle \Omega, s \rangle$ is a Similarity-Projection (SP) structure as defined in [5], where p was used instead of s. Intuitively, SP-structures may be understood as one-dimensional subspaces of a Hilbert space with holes. A set of n elements is an n-dimensional Hilbert space with very big holes. We shall now recall the properties of SP-structures that we shall need, with the necessary definitions and notations. We restrict our attention to *standard* SP-structures as defined in [5]. The definition of a standard SP-structure is recalled in Appendix A.

The properties below are the ones we shall use in the sequel, they should not be taken as a definition of SP-structures. A physically and epistemologically motivated definition of SP-structures may be found in [5] where the properties below are proved out of a set of seemingly weak assumptions. Property 7 that is so striking in Hilbert spaces is not an assumption, it follows from more basic properties. Similarly for Property 8. Property 13 seems original. It means that the similarity function $s(x, y)$ is, in a sense, continuous: for $\epsilon > 0$, close enough to 0, if $s(x, y) \geq 1 - \epsilon$, then for any $z \in \Omega$ the difference $s(x, z) - s(y, z)$ is of order $\sqrt{\epsilon}$.

In the following x, y, z are arbitrary elements of the sample space Ω and A, B are arbitrary subsets of Ω.

1. $s(x, y) \in [0, 1]$, and $x = y$ iff $s(x, y) = 1$,
2. $s(y, x) = s(x, y)$,
3. x and y are said to be *orthogonal*, written $x \perp y$ iff $s(x, y) = 0$, we say that x is orthogonal to A and write

$x \perp A$ iff $x \perp y$ for every $y \in A$, we say that A and B are orthogonal and write $A \perp B$ iff $z \perp B$ for every $z \in A$,

4. A is said to be an *ortho-set* iff all pairs of distinct elements of A are orthogonal,

5. for any ortho-set A, $s(x, A) \stackrel{\text{def}}{=} \sum_{y \in A} s(x, y) \leq 1$,

6. B is said to be a *subspace* and A is said to be a basis for B iff A is an ortho-set and $B = \{x \in \Omega \mid s(x, A) = 1\}$,

7. if B is a subspace all bases for B have the same cardinality,

8. if A_i for $i \in I$ are subspaces, then their intersection $\bigcap_{i \in I} A_i$ is also a subspace: subspaces are closed under arbitrary intersections,

9. $\emptyset$ is a subspace, Ω is a subspace,

10. the orthogonal complement of any subset A is defined by:
$$A^{\perp} \stackrel{\text{def}}{=} \{x \in \Omega \mid x \perp A\},$$

11. $A^{\perp}$ is a subspace, if A is a subspace then $(A^{\perp})^{\perp} = A$, $\emptyset^{\perp} = \Omega$, $\Omega^{\perp} = \emptyset$,

12. for any subspace A and any $x \in \Omega$, if x is not orthogonal to A, there is a unique $t(x, A) \in A$ such that $s(x, t(x, A)) = s(x, A)$ and for every $y \in A$ one has $s(x, y) = s(x, t(x, A))\, s(t(x, A), y)$,

13. $$s(z, x) \leq \tag{1}$$
$$s(z, y) \; + \; 1/2 \sqrt{1 - s(x, y)} \; + \; (1 - s(x, y)).$$

Note that the seemingly natural *triangular inequality*: $s(x, y) \leq s(x, z)\, s(z, y)$ is not a property of SP-structures. It does not hold in Hilbert spaces. A classical SP-structure is defined to be a structure in which $s(x, y) = 0$ whenever $x \neq y$. In a classical SP-structure x and y are orthogonal iff they are different, and A and B are orthogonal iff they are disjoint. Any set A is a subspace. The orthogonal complement of a set A is its set complement: $\Omega - A$.

5 Properties of SP structures

We present here properties of SP-structures that have not been presented in [5]. We define the sum $A \oplus B$ of any two *subsets* of Ω. The set $A \oplus B$ is the minimal *subspace* that contains A and B.

Definition 4 *Let $\langle \Omega, s \rangle$ be an SP-structure. If $A, B \subseteq \Omega$, their sum $A \oplus B$ is defined to be the smallest subspace including $A \cup B$:*

$$A \oplus B \;=\; \bigcap_{X \text{ is a subspace, } A \cup B \subseteq X} X.$$

This definition is correct since, as noticed in 8 above, subspaces are closed under intersection. One easily sees that sum is commutative, associative and monotone: $A \subseteq A'$ implies $A \oplus B \subseteq A' \oplus B$. Therefore the sum of any family (finite or infinite) of subsets is well-defined: $\bigoplus_{i \in I} A_i$ is the intersection of all subspaces including $\cup_{i \in I} A_i$.

In a classical structure sum is union: $A \oplus B = A \cup B$.

Lemma 1 *For any subspaces A, B: $A \oplus A^{\perp} = \Omega$ and $A \cap A^{\perp} = \emptyset$.*

Proof: Let $x \in \Omega$. Let B be a basis for A. By Theorem 1 of [5] there is a basis for Ω that includes B. Let $B \cup B'$ be this basis: $p(x, B) + p(x, B') = 1$ and $B' \subseteq A^{\perp}$. Any subspace that includes B and B' must be Ω. We have shown that $A \oplus A^{\perp} = \Omega$.

If $x \in A \cap A_{\perp}$ we must have $s(x, x) = 0$, contradicting property 1 above. We have shown that $A \cap B = \emptyset$. ▮

We shall now show that the set of subspaces of an SP-structure is an orthomodular complemented lattice. The lattice of closed subspaces of a Hilbert space shows that it is not always modular. The structure we are interested in is an orthomodular lattice, but note that we have additional structure given by the similarity function.

Theorem 1 *Let $\langle \Omega, s \rangle$ be an SP-structure. The set of subspaces of Ω is a complete complemented orthomodular lattice, if one takes $A \leq B$ iff $A \subseteq B$. Least upper bound is $\oplus$ and greatest lower bound is intersection.*

Proof: The relation $\leq$ is obviously a partial order and, since, by 8, subspaces are closed under intersections, intersections are greatest lower bounds. By definition sums are least upper bounds and the lattice is complete. Lemma 1 shows that it is a complemented lattice. Orthomodularity is a consequence of Theorem 8 of [5]: if $A \subseteq C$ are subspaces any basis B for A can be extended into a basis $B \cup B'$ for C and therefore $C \subseteq B \oplus B' \subseteq A \oplus A^{\perp} \cap C$. ▮

De Morgan's laws hold in any orthocomplemented lattice.

Corollary 1 *For any subspaces A and B $(A \cap B)^{\perp} = A^{\perp} \oplus B^{\perp}$ and $(A \oplus B)^{\perp} = A^{\perp} \cap B^{\perp}$. These equalities extend to arbitrary infinite sums and intersections.*

We shall now generalize the similarity s to *arbitrary subspaces* of Ω.

Definition 5 *Let $A, B \subseteq \Omega$ be subspaces. We wish to define a measure of their similarity, denoted $s(A, B)$. Let $x \in \Omega$, we shall define $\tau(x, A, B)$ to be the similarity of A and B from the vantage point x. Then we let $s(A, B) = \liminf\{\tau(x, A, B) \mid x \in \Omega\}$. Now let us define $\tau(x, A, B)$. In case $x \not\perp A$ and $x \not\perp B$, let $\tau(x, A, B) = s(t(x, A), t(x, B))$. If $x \perp A$, we let $\tau(x, A, B) = 1 - s(x, B)$. If $x \perp B$, we let $\tau(x, A, B) = 1 - s(x, A)$.*

Note that if $x \perp A$ and $x \perp B$ both last conditions give $\tau(x, A, B) = 1$. Note also that $s(A, B) = s(B, A)$.

Theorem 2 *For any $x, y \in \Omega$, we have $s(\{x\}, \{y\}) = s(x, y)$.*

Proof: If $z \not\perp x$ and $z \not\perp y$ we have $\tau(z, \{x\}, \{y\}) = s(x, y)$. If $z \perp x$, we have $\tau(z, x, y) = 1 - s(z, y) \geq s(x, y)$. If $z \perp y$, we have $\tau(z, x, y) = 1 - s(z, x) \geq s(x, y)$. ∎

Theorem 3 *If $x \in A$, then $s(A, B) \leq s(x, B)$.*

Proof:

Suppose $x \not\perp B$. Then one has $s(A, B) \leq \tau(x, A, B) = s(x, t(x, B)) = s(x, B)$. If $x \perp B$ we have $s(A, B) \leq \tau(x, A, B) = 1 - s(x, A) = 0 \leq s(x, B)$.

∎

In general, $s(\{x\}, B) < s(x, B)$.

Theorem 4 *$s(A, B) = 1$ iff $A = B$.*

Proof: Let $A = B$. If $x \not\perp A$ we have $\tau(x, A, B) = s(t(x, A), t(x, A)) = 1$. If $x \perp A$ we have $\tau(x, A, B) = 1 - s(x, B) = 1$. Assume, now, that $s(A, B) = 1$. For every $x \in \Omega$ we have $\tau(x, A, B) = 1$. Suppose $x \in A$. If $x \not\perp B$ we have $\tau(x, A, B) = s(x, t(x, B)) = 1$ and $x \in B$. If $x \perp B$, we have $1 - s(x, t(x, A)) = 1$ and $s(x, x) = 0$, a contradiction. We conclude that $A \subseteq B$. Similarly we have $B \subseteq A$ and we conclude that $A = B$. ∎

We may now generalize Inequality 1.

Theorem 5 *Let $A, B, C \subseteq \Omega$ be subspaces. We have*

$$s(A, B) \leq s(A, C) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C).$$

Proof: It is enough to show that, for any $x \in \Omega$, we have

$$\tau(x, A, B) \leq \tag{2}$$

$$\tau(x, A, C) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C).$$

Suppose, first, that $x \not\perp A$, $x \not\perp B$ and $x \not\perp C$. Let $w = t(x, A)$, $y = t(x, B)$ and $z = t(x, C)$. By Inequality 1 we have:

$$s(w, y) \leq s(w, z) + 1/2\sqrt{1 - s(y, z)} + 1 - s(y, z)$$

and therefore

$$\tau(x, A, B) \leq$$

$$\tau(x, A, C) + 1/2\sqrt{1 - \tau(x, B, C)} + 1 - \tau(x, B, C).$$

But $s(B, C) \leq \tau(x, B, C)$ and Equation 2 is proved. Suppose, now, that $x \not\perp A$, $x \not\perp B$ but $x \perp C$. Let $w = t(x, A)$ and $y = t(x, B)$. We must show that

$$s(w, y) \leq 1 - s(x, A) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C).$$

But $s(x, A) = s(x, w)$. We know that $x \perp C$ and $s(w, x) + s(w, C) \leq 1$. It is therefore enough to show that

$$s(w, y) \leq s(w, C) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C).$$

If $w \not\perp C$, by Inequality 1, we have

$$s(w, y) \leq$$

$$s(w, t(w, C)) + 1/2\sqrt{1 - s(\{y\}, C)} + 1 - s(\{y\}, C).$$

We conclude by Theorem 3. If $w \perp C$, it is enough to show that $s(w, y) + s(B, C) \leq 1$. But $s(B, C) \leq s(y, C)$ by Theorem 3 and we have $s(y, w) + s(y, C) \leq 1$.

The case $x \not\perp A$, $x \not\perp C$ but $x \perp B$ is treated similarly.

If $x \not\perp A$, $x \perp B$ and $x \perp C$ we must show that

$$1 - s(x, A) \leq 1 - s(x, A) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C)$$

which is obvious.

We are left with the case $x \perp A$. We must show that

$$1 - s(x, B) \leq 1 - s(x, C) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C),$$

or equivalently

$$s(x, C) \leq s(x, B) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C).$$

If $x \perp C$ the claim is obvious. Assume $x \not\perp C$ and let $z = t(x, C)$. If $x \perp B$ we have $s(z, x) + s(z, B) \leq 1$ and therefore $s(x, C) \leq 1 - s(z, B) \leq 1 - s(B, C)$. If, last, $x \not\perp B$ and we let $y = t(x, B)$, by Inequality 1 we have:

$$s(x, z) \leq s(x, y) + 1/2\sqrt{1 - s(y, z)} + 1 - s(y, z)$$

and therefore

$$s(x, C) \leq s(x, B) + 1/2\sqrt{1 - \tau(x, B, C)} + 1 - \tau(x, B, C) \leq$$

$$s(x, B) + 1/2\sqrt{1 - s(B, C)} + 1 - s(B, C).$$

∎

In classical SP-structures $s(A, B)$ is equal to 1 iff $A = B$ and equal to 0 otherwise.

6 σ*-fields

We want to generalize Definition 1, i.e., the definition of a σ-field on a set Ω to that of a σ*-*field* over an SP-structure $\langle \Omega, s \rangle$. As expected, we require that the σ*-field $\mathcal{F}$ be closed under countably many sums (sums generalize unions) and under orthogonal complement (they generalize set-complements). But we require the elements of a σ*-field $\mathcal{F}$ are *subspaces*, not arbitrary subsets, of Ω.

Definition 6 *Let $\langle \Omega, s \rangle$ be an SP-structure. A set $\mathcal{F}$ of* subspaces *of $\langle \Omega, s \rangle$is said to be a σ**-field *over $\langle \Omega, s \rangle$ iff:*

1. *$\emptyset \in \mathcal{F}$,*
2. *for every $A \in \mathcal{F}$, its orthogonal complement $A^{\perp}$ is in $\mathcal{F}$,*
3. *for any set, finite or* countably *infinite A_i, $i \in I$, of pairwise orthogonal elements of $\mathcal{F}$, its sum $\bigoplus_{i \in I} A_i$ is in $\mathcal{F}$.*

Elements of $\mathcal{F}$ are called events.

Note that $\Omega = \emptyset^{\perp}$ is an event.

If $\langle \Omega, s \rangle$ is a classical SP-structure then the notion of a σ*-field on the structure is equivalent to that of a σ-field on Ω.

Lemma 2 *Assume $\mathcal{F}$ is a σ*-field on $\langle \Omega, s \rangle$, I is finite or* countably *infinite, and for any $i \in I$, A_i is an element of $\mathcal{F}$. Then, the intersection $\bigcap_{i \in I} A_i$ is in $\mathcal{F}$.*

Proof: By Corollary 1. ∎

Corollary 2 *Any σ*-field is a bounded complemented orthomodular lattice, if one takes $A \leq B$ iff $A \subseteq B$. Least upper bound is $\oplus$ and greatest lower bound is intersection. Countably infinite sets have l.u.b. and g.l.b. but the lattice is not, in general, complete.*

7 Probability distributions

We may now generalize Definition 2. We shall define *-probabilities that attach a probability to events of a σ*-field. Note that *states* in Quantum Physics are such *-probabilities. Our first three conditions are those of Definition 2, but a fourth condition is added to ensure that probabilities are, in a sense, *continuous*. If the subspaces A and B are *close*, i.e., $s(A, B)$ is close to 1, then we expect $p(A)$ and $p(B)$ to be close to each other.

Definition 7 *Assume $\langle \Omega, s \rangle$ is an SP-structure, and $\mathcal{F}$ is a σ-* field on $\langle \Omega, s \rangle$. A *-probability on $\langle \Omega, s, \mathcal{F} \rangle$ is a function $p : \mathcal{F} \longrightarrow [0, +\infty]$ that satisfies:*

1. *$p(\emptyset) = 0$,*
2. *$p(\Omega) = 1$,*
3. *for any finite or countably infinite set, A_i, $i \in I$ of pairwise orthogonal elements of $\mathcal{F}$ one has: $p(\bigoplus_{i \in I} A_i) = \sum_{i \in I} p(A_i)$,*
4. *for any events A, B, we have*

$$p(A) \leq \tag{3}$$

$$p(B) + 1/2\, \sqrt{1 - s(A, B)} + (1 - s(A, B)).$$

Note that, in our third condition, the sum $\bigoplus_{i \in I} A_i$ is an event by Definition 6. Our fourth condition is taken from 13 above, which has been shown to be tight in [5].

It is clear that convex combinations of probabilities are probabilities.

Lemma 3 *Assume $\langle \Omega, s \rangle$ and $\mathcal{F}$ are fixed. If for any $i \in I$ p_i is a probability and $w_i \in [0, 1]$ are such that $\sum_{i \in I} w_i = 1$, then $q = \sum_{i \in I} w_i\, p_i$ defined by $q(A) = \sum_{i \in I} w_i\, p_i(A)$ for any $A \in \mathcal{F}$ is a probability.*

8 Pure and mixed states

In Quantum Physics pure states have a dual aspect: they are points of the sample space, i.e., elements of the subspaces representing quantic propositions, but they also attach probabilities to points and subspaces (the *transition probability*). This simply generalizes the fact that a point x in the sample space can be identified, in Kolmogorov's setting, with the probability distribution that gives probability one to all events that contain x and probability zero to all other events. Probabilities attached to points in the sample space are called *pure states* in Quantum Physics.

Theorem 6 *Assume $\langle \Omega, s \rangle$ is an SP-structure, and $\mathcal{F}$ is a σ-* field on $\langle \Omega, s \rangle$. Let $x \in \Omega$ be a point in the sample space. One may define a *-probability p_x by: $p_x(B) = s(x, B) = \sum_{y \in A} s(x, y)$ for any event B and any basis A for B. Such probabilities are called* pure states *and the set of pure states will be denoted by $P(\Omega)$. Convex combinations of pure states are called mixed states. The set of mixed states will be denoted $M(\Omega)$. We shall represent mixed states as convex combinations of points of the sample space: $p = \sum_{i \in I} r_i\, x_i$ for non-negative real numbers r_i such that $\sum_{i \in I} r_i = 1$ and $x_i \in \Omega$ for $i \in I$.*

Proof: Obviously $p_x(\emptyset) = 0$ and $p_x(\Omega) = 1$. Suppose now that B_i, $i \in I$ is a family of pairwise orthogonal events. We have $s(x, \bigoplus_{i \in I} B_i) = \sum_{i \in I} s(x, B_i)$ since a basis for the sum is the union of bases for the B's. To check the last (continuity) property of probability measures, assume, first, that x is orthogonal to neither A nor B. By properties 12 and 13 of Section 4 and by Definition 5 we have:

$$s(x, A) = s(x, t(x, A)) \leq$$
$$s(x, t(x, B)) + 1/2\sqrt{1 - s(t(x, A), t(x, B))} +$$
$$(1 - s(t(x, A), t(x, B))) \leq$$
$$s(x, B) + 1/2\sqrt{1 - s(A, B)} + (1 - s(A, B)).$$

If $x \perp A$, the claim is obvious. Suppose, now that $x \perp B$ and $x \not\perp A$. We have $s(A, B) \leq \tau(x, A, B) = 1 - s(x, A)$. Therefore $s(x, A) \leq 1 - s(A, B)$. ∎

Gleason's theorem [6] says that, for any SP-structure defined by the rays of a Hilbert space of dimension larger than 2, any probability measure on the σ*-field of all closed subspaces is a mixed state. Notice that the result does not hold for Hilbert spaces of dimension 2. For physical systems of dimension 2 there are probabilities that are not mixed states. Nevertheless it seems that the only probabilities found useful to study such systems in quantum physics are mixed states. The reason may be hidden in the preparation of quantic systems: one seems to know how to prepare a system in any mixed state but not in any state corresponding to a probability measure that is not mixed. Therefore one is probably justified in restricting one's attention to mixed states.

A most important remark is that the set $M(\Omega)$ of all convex combination of pure states is not a free structure. We may well have, for example $1/2\, p_x + 1/2\, p_y = 1/2\, p_w + 1/2\, p_z$ with $x \neq w$ and $x \neq z$. A topic for further study is the characterization of those transformations $\tau : P(\Omega) \longrightarrow M(\Omega)$ for which $\sum_{i \in I} r_i p_{x_i} = \sum_{j \in J} s_j p_{y_j}$ implies $\sum_{i \in I} r_i \tau(p_{x_i}) = \sum_{j \in J} s_j \tau(p_{y_j})$.

In classical structures, mixed states are discrete probability measures and therefore the remainder of this paper generalizes only discrete probability theory. A generalization of continuous probability theory is probably necessary to understand systems with observables that can take a continuum of values.

9 Random variables

The definition of *-random variables, generalizing Definition 3 requires some thinking.

Definition 8 *Let $\langle \Omega_i, s_i \rangle$ be SP-structures, and $\mathcal{F}_i$ be σ-* fields on $\langle \Omega_i, s_i \rangle$ for $i = 1, 2$. We want a random variable to give values in Ω_2 to elements of Ω_1. So it seems a random variable X should be a function $\Omega_1 \longrightarrow \Omega_2$. But we noticed in Section 1 that non-commuting observables cannot be defined have values at the same sample points. Therefore we must accept the idea that X be a* partial *function $X : \Omega_1 \longrightarrow \Omega_2$. In the classical case of Definition 3, the function is a* total *function and therefore we shall require that X be defined on some basis for Ω_1. In the classical case Ω_1 is the only basis and therefore X must be total. We, then, as usual, require that the inverse image by X of any element of $\mathcal{F}_2$ be an element of $\mathcal{F}_1$. Guided by the fact that, in the classical case, if A, B are* disjoint *elements of Ω_2, their inverse images $X^{-1}(A)$ and $X^{-1}(B)$ are disjoint, we require that if $A, B \in \mathcal{F}_2$ and $A \perp B$, we have $X^{-1}(A) \perp X^{-1}(B)$.*

Real random variables are important enough to justify a specialization of Definition 9

Definition 9 *Let $\langle \Omega, s \rangle$ be an SP-structure, and $\mathcal{F}$ a σ-* field on $\langle \Omega, s \rangle$. A real random variable X is a partial function $X : \Omega \longrightarrow \mathcal{R}$ that is defined on some basis for Ω and such that the inverse image by X of any Lebesgue-measurable subset of $\mathcal{R}$ is an element of $\mathcal{F}$ and such that the inverse images of any two disjoint such subsets are orthogonal elements of $\mathcal{F}$.*

Note that Definition 9 ensures that the set of points of the sample space Ω on which a random variable X is defined is a set of pairwise orthogonal subspaces (generalizing eigen-subspaces) whose sum is Ω.

A real random variable is a partial function, but it defines a total function: its expected value in each state. There is no problem in considering that expected values of non-commuting observables are both defined at the same time. This total function can be even defined on mixed states.

Definition 10 *Let X be a real random variable as above and suppose it takes only a countable set of values: r_i for $i \in I$. Let $p \in M(\Omega)$ be any mixed state. We define $\hat{X}(p)$ as $\sum_{i \in I} r_i\, p(X^{-1}(r_i))$.*

Theorem 7 *Let X be a random variable as in Definition 10. Let $x \in \Omega$ and assume $B = \{b_i \mid i \in I\}$ is a basis for Ω on which X is defined. Then $\hat{X}(p_x) = \sum_{i \in I} X(i)\, s(x, b_i)$.*

Proof: For any $a \in \mathcal{R}$, let $J(a) \subseteq I$ be the set of indexes i for which $X(i) = a$. The subspace $\oplus_{i \in J} b_i$ spanned by the corresponding basis elements is equal to $X^{-1}(a)$, and $p_x(X^{-1}(a)) = s(x, \oplus_{i \in J} b_i) = \sum_{i \in J} s(x, b_i)$. ∎

10 Future work

In Quantum Physics, operators, and particularly self-adjoint operators, play a central role. Operators can be composed and their commutation properties represent important physical information. One should try to reflect this *transformational* aspect into our present framework, in terms of properties of *-random variables. We hope to be able to characterize classical Kolmogorov's probability theory as the special case of *-probabilities in which random variables commute.

Acknowledgements

I want to thank Itamar Pitowsky: a lecture of his started me thinking on the relation of my previous work to Quantum Probability and his help and advice are gratefully acknowledged and Benjamin Weiss for his help.

A SP-structures

We recall here the essentials of the definition of SP-structures as presented in [5], with minimal explanations. The reader is referred to [5] for a gentle complete introduction.

Property 1 (Symmetry) *For any $x, y \in \Omega$, $s(y, x) = s(x, y)$.*

Symmetry is an experimentally verifiable and fundamental property of Quantum Mechanics, see, e.g., the Law of Reciprocity in [7], p. 35.

Property 2 (Non-negativity) *For any $x, y \in \Omega$, $s(x, y) \geq 0$.*

Property 3 (Boundedness) *For any state $x \in \Omega$ and any ortho-set A, $s(x, A) \stackrel{\text{def}}{=} \sum_{a \in A} s(x, a) \leq 1$.*

The next property we want to consider deals with orthogonal projections.

Property 4 (O-Projection) *Suppose $x \in \Omega$ is a state and $A \subseteq \Omega$ is an ortho-set such that $s(x, A) < 1$. Then there exists a state $y \in \Omega$ with the following properties:*

1. *$y \perp A$, i.e., $s(y, A) = 0$, i.e., $A \cup \{y\}$ is an ortho-set, and*
2. *$s(x, A) + s(x, y) = 1$.*

O-Projection should remind the reader of the Gram-Schmidt process. Physically, the ortho-set A represents certain values of a given observable and therefore can be interpreted as a test: is the state x in A or not. If $s(x, A) < 1$ the answer to the question above may, with a certain "probability" be "no". If the answer is indeed "no" the system is left in a state y that satisfies the three conditions above. The scalar product can be seen to satisfy those conditions, when y is the projection of x on the subspace $A^{\perp}$ orthogonal to A. In a classical system, $s(x, A) < 1$ implies $s(x, A) = 0$ and we can take $y = x$.

Definition 11 *If A is an ortho-set, the subspace $\bar{A} \subseteq \Omega$ generated by A is defined by: $\bar{A} = \{x \in \Omega \mid s(x, A) = 1\}$. The ortho-set A is said to be a basis for $\bar{A}$.* A basis *is a basis for Ω. A subspace is a set of states $X \subseteq \Omega$ such that there exists some ortho-set A such that $X = \bar{A}$.*

Our next defining property for SP-structure is a factorization property.

Property 5 (Factorization) *Let A be an ortho-set and x an arbitrary state. If $y, z \in \bar{A}$ and $s(x, y) = s(x, A)$, then $s(x, z) = s(x, y)\, s(y, z)$.*

Factorization implies that $s(x, A)$ is the maximum of all $s(x, y)$ for $y \in \bar{A}$ and that every such $s(x, y)$ can be factored out through the state taking this maximum. Factorization has been described in Theorem 1 of [4]. The meaning of Factorization, for Physics, is that, if one knows that in state y some observable A has a specific value, then the probability of a transition from x to y is the product of the probability of measuring this specific value (in x) times the transition probability from the state obtained after the measurement to y. Factorization seems to be a logical requirement relating tests to two propositions one of which entails the other: if A entails B, testing for A may be done by testing first for B and then for A.

In [5] a last property is presented that is shown to imply Equation 1 of Section 4. Since, in this paper, we only need Equation 1, we shall not present this property here.

Definition 12 *Any two states $x, y \in \Omega$ are said to be* equivalent, *and we write $x \sim y$ iff for any $z \in \Omega$, one has: $s(x, z) = s(y, z)$. An SP-structure is said to be* standard *if any two equivalent states are equal.*

References

[1] Garret Birkhoff and John von Neumann. The logic of quantum mechanics. *Annals of Mathematics*, 37:823–843, 1936.

[2] Maria Luisa Dalla Chiara. Quantum logic. In D. M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 6, pages 129–228. Kluwer, Dordrecht, 2nd edition edition, 2001.

[3] Daniel Lehmann. A presentation of quantum logic based on an "and then" connective. *Journal of Logic and Computation*, 18(1):59–76, February 2008. DOI: 10.1093/logcom/exm054.

[4] Daniel Lehmann. Quantic superpositions and the geometry of complex Hilbert spaces. *International Journal of Theoretical Physics*, 47(5):1333–1353, May 2008. DOI:10.1007/s10773-007-9576-y.

[5] Daniel Lehmann. Similarity-projection structures: the logical geometry of quantum physics. *International*

Journal of Theoretical Physics, 48(1):261–281, 2009. DOI: 10.1007/s10773-008-9801-3.

[6] Gleason A. M. Measures on the closed subspaces of a hilbert space. *Journal of Mathematics and Mechanics*, 6:885–893, 1957.

[7] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer, Dordrecht, The Netherlands, 1995.

[8] Itamar Pitowsky. Quantum mechanics as a theory of probability. In W. Demopoulos and I. Pitowsky, editors, *Festschrift in Honor of Jeffrey Bub*. forthcoming.