Generalized probabilities in statistical theories

Federico Holik,[1,2] C. Massri, A. Plastino[1,3] and Manuel Sáenz

June 4, 2014

1- Universidad Nacional de La Plata, Instituto de Física (IFLP-CCT-CONICET), C.C. 727, 1900 La Plata, Argentina
2- Departamento de Matemática - Ciclo Básico Común
Universidad de Buenos Aires - Pabellón III, Ciudad Universitaria
Buenos Aires, Argentina
3- Center Leo Apostel for Interdisciplinary Studies and, Department of Mathematics, Brussels Free University Krijgskundestraat 33, 1160 Brussels, Belgium

Abstract

In this review article we present different formal frameworks for the description of generalized probabilities in statistical theories. We discuss the particular cases of probabilities appearing in classical and quantum mechanics, possible generalizations of the approaches of A. N. Kolmogorov and R. T. Cox to non-commutative models, and the approach to generalized probabilities based on convex sets.

Key words:Quantum Probability; Lattice theory; Information theory; Classical probability; Cox's approach

1 Introduction

Which is the nature of the probabilities appearing in Quantum Mechanics (QM)? The Born's rule tell us how to compute the probability of a given process using the formalism of QM. But is it legitimate to speak of something like Quantum Probabilities (QP)? Some authors react against this possibility, arguing that this is not possible, mainly because of the non-Kolmogorovian character of the probabilities appearing in QM. Their central point is that probabilities in QM cannot be described using classical probability theory because they do not obey Kolmogorov's axioms.

On the other hand, many authors support the idea that a quantum probability theory is possible by generalizing Kolmogorov's axioms to non boolean algebras [76, 26, 78, 61, 35, 60]. In this sense, QP exists as a non-Boolean (or non-commutative) generalization of classical probability theory.

However, Kolmogorov's approach is not the only one. An alternative approach to classical probability theory is presented by R. T. Cox [23, 22]. It is based on a study of the measure functions compatible with the algebraic properties of the logic of an intelligent agent trying to make probable inferences out of available data. A variant of this approach has been used

to describe probabilities in QM [24, 36, 52, 53, 54, 55, 37, 56, 44], suggesting that they are essentially classical.
We have already mentioned that there exists a general formal framework which includes CP and QP as particular cases, being a generalization of Kolmogorov's approach to non-distributive algebras. Is there an analogue of this for the Cox approach? An affirmative answer to this question was given in [44]. It is shown there that the typical features of QP arise whenever the lattice of propositions of the Cox's approach is replaced by a non-distributive one. As is well known, the quantum-logical approach to QM characterizes this theory using a lattice-theoretical approach in which the lattice is an orthomodular one [7, 27, 77, 8, 85, 86, 79, 39, 45, 69, 50, 51, 33, 34, 70, 29, 2, 3, 4, 5, 6]. In [44] it is shown that when the Cox's method is applied to the orthomodular lattice of projections of a Hilbert space, QP are derived.
Remark that generalized probabilities can also be studied in what has been called the Convex Operational Models (COM) approach [14, 90, 9, 10, 11, 12, 13, 80, 15, 71, 16, 43]. In this approach, the properties of the systems studied and their associated probabilities are encoded in a geometrical way in the form of a convex set and its space of observable magnitudes. The quantum formalism and many aspects of quantum information theory (such as entanglement, discord, information protocols) can be suitably described using this approach [13, 80, 15, 12, 71, 43, 16, 11, 9, 10, 71]. Non-linear generalizations of QM where studied using the convex approach in [62, 63, 64].

The present is a review article and no new technical results are presented. It can be considered as an extension of [38], in which, due to reasons of space, we could not discuss in detail some interesting topics. This notwithstanding, the presentation of the topics is original in the sense that the perspective presented has not been previously published in the literature, and has conceptual implications for the discussion about what is the meaning of CP and QP. We start by reviewing lattice theory and the quantum logical approach to QM in Sections 2 and 3. Next, in Section 4, we concentrate on the different approaches to CP, namely, Kolmogorov's and Cox's, and in Section 5 we review the COM approach. After these notions are introduced, we discuss QP in Section 6. We present the generalization of the Cox's method to general non-distributive lattices in Section 7. Finally, some conclusions are drawn in Section 8.

2 Lattice Theory

Lattices can be defined by using equations, i.e., they can be characterized as algebraic structures satisfying certain axiomatic identities. A set $\mathcal{L}$ endowed with two operations $\wedge$ and $\vee$ will be called a *lattice*, if for al $x, y, z \in \mathcal{L}$ the following equations are satisfied

$$x \vee x = x \;\; x \wedge x = x \;\; \text{(idempotence)} \tag{1a}$$

$$x \vee y = y \vee x \;\; x \wedge y = y \wedge x \;\; \text{(commutativity)} \tag{1b}$$

$$x \vee (y \vee z) = (x \vee y) \vee z \;\; x \wedge (y \wedge z) = (x \wedge y) \wedge z \;\; \text{(associativity)} \tag{1c}$$

$$x \vee (x \wedge y) = x \wedge (x \vee y) = x \;\; \text{(absortion)} \tag{1d}$$

If the extra relationships

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \;\; \text{(distributivity 1)} \tag{2a}$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \;\; \text{(distributivity 2)} \tag{2b}$$

are satisfied, the lattice is called *distributive*.

Lattice theory can also be studied using *partially ordered sets* (*poset*). A poset is a set X endowed with a partial ordering relation "$<$" satisfying

- For all $x, y \in X$, if $x < y$ and $y < x$, then $x = y$
- For all $x, y, z \in X$, if $x < y$ and $y < z$, then $x < z$

We use the notation "$x \leq y$" for the case "$x < y$" or "$x = y$". A lattice $\mathcal{L}$ will be a poset for which any two elements x and y have a unique supremum and a unique infimum with respect to the order structure. The least upper bound of two given elements "$x \vee y$" is called the "join" and their greatest lower bound "$x \wedge y$", called their "meet". A lattice for which all its subsets have both a supremum and an infimum is called a *complete lattice*. If furthermore there exists a greatest element $\mathbf{1}$ and a least element $\mathbf{0}$, the lattice is called *bounded*. They are usually called the *maximum* and the *minimum* respectively. Any lattice can be extended into a bounded lattice by adding a greatest and a least element. Every non-empty finite lattice is bounded. Complete lattices are always bounded. An orthocomplementation in a bounded poset P is a unary operation "$\neg(\ldots)$" satisfying:

$$\neg(\neg(a)) = a \tag{3a}$$

$$a \leq b \longrightarrow \neg b \leq \neg a \tag{3b}$$

$a \vee \neg a$ and $a \wedge \neg a$ exist and

$$a \vee \neg a = \mathbf{1} \tag{3c}$$

$$a \wedge \neg a = \mathbf{0} \tag{3d}$$

hold.
A bounded poset with orthocomplementation will be called an *orthoposet*. An *ortholattice*, will be an orthoposet which is also a lattice. For a, $b \in \mathcal{L}$ (an ortholattice or orthoposet), we say that a is orthogonal to b ($a \perp b$) iff $a \leq \neg b$. Following [74], we define an *orthomodular lattice* as an ortholattice satisfying the orthomodular law:

$$a \leq b \text{ and } \neg a \leq c \Longrightarrow a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \tag{4}$$

A *modular lattice*, is an ortholattice satisfying the stronger condition (modular law)

$$a \leq b \Longrightarrow a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \tag{5}$$

and finally, a *boolean lattice* will be an ortholattice satisfying the still stronger condition (distributive law)

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \tag{6}$$

If $\mathcal{L}$ has a null element 0, then an element x of $\mathcal{L}$ is an *atom* if $0 < x$ and there exists no element y of $\mathcal{L}$ such that $0 < y < x$. $\mathcal{L}$ will be said to be

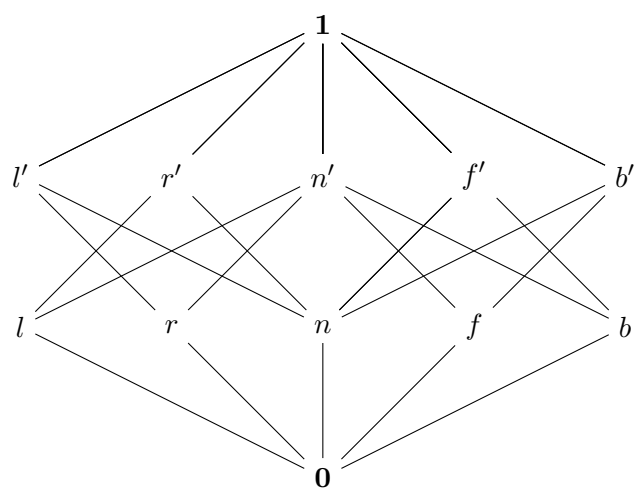
- *Atomic*, if for every nonzero element x of $\mathcal{L}$, there exists an atom a of $\mathcal{L}$ such that $a \leq x$.
- Atomistic, if every element of $\mathcal{L}$ is a supremum of atoms.

In order to illustrate the above notions, let us consider the following examples:

Hilbert lattice: As an example of an orthomodular lattice which will be of importance for us, let us consider the set of closed subspaces of a Hilbert space $\mathcal{H}$. They can be endowed with a lattice structure as follows [77]. The operation"$\vee$" is taken as the direct sum "$\oplus$" of subspaces, "$\wedge$" as the intersection "$\cap$", and "$\neg$" as the orthogonal complement "$\perp$", $\mathbf{0} = \vec{0}$, $\mathbf{1} = \mathcal{H}$, and denote by $\mathcal{P}(\mathcal{H})$ to the set of closed subspaces. Then, $< \mathcal{P}(\mathcal{H}),\ \cap,\ \oplus,\ \neg,\ 0,\ 1 >$ will be a complete bounded orthomodular lattice (which we will denote simply by $\mathcal{P}(\mathcal{H})$). As closed subspaces are in one to one correspondence with projection operators, we will take $\mathcal{P}(\mathcal{H})$ to be the lattice of closed subspaces or the lattice of projections interchangeably. One of the most important features of $\mathcal{P}(\mathcal{H})$ is that the distributive law (6) doesn't holds. $\mathcal{P}(\mathcal{H})$ is modular iff $\mathcal{H}$ is finite dimensional. If $\mathcal{H}$ is infinite dimensional, then $\mathcal{P}(\mathcal{H})$ is always orthomodular.

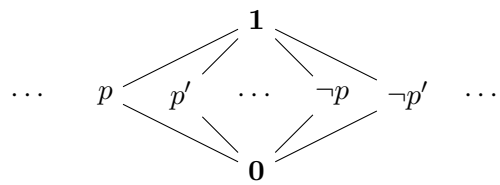
Subsets of a given set As an example of a boolean lattice, consider the subsets of a given set endowed with the operations union "$\cup$" as "$\vee$", "$\cap$" as "$\wedge$", and set theoretical complement as "$\neg$". The propositional calculus of classical logic also forms a boolean lattice.

Chinesse lantern: The (chinesse lantern)[79] is used in quantum logic to show an example of a system which is not a quantum one neither a classical one. Its lattice of propositions corresponds to the following Hasse diagram



looking at the above diagram, it is easy to see that this lattice contains two Boolean sublattices.

The lattice of Q-bit: Given the incredible advances of quantum information theory in the last decades, the reader may wonder how does the lattice of a q-bit looks like. Suppose then that we are given a spin $\frac{1}{2}$ system. As is well known, the set of all possible pure states is isomorphic to a sphere, namely, the *Bloch sphere* [20]. The different sets of objective properties, which are of the form "the particle has spin + (or −) in direction $\hat{z}$", are represented by rays (or points on the surface of the sphere). They can be given by $\mathbf{0}$, the null element of the space, $\mathbf{1}$ (the maximal element of the space) or a ray. Each ray will have associated a projection operator $|+\rangle_{\hat{z}}\langle+|$. Thus, the Hasse diagram of a q-bit will have the form:



3 The Formalism of QM

In this Section we discuss some specific features of the quantum formalism [87, 85, 86, 77] which are relevant for the problem of QP.

3.1 Elementary measurements and projection operators

The rigorous formulation of the formalism of QM was presented in a series of papers by von Neumann, Jordan, Hilbert and Nordheim [58]. It can be said that its definitive form was accomplished in the book of von Neumann [87]. In QM, observable physical magnitudes are represented by compact self-adjoint operators in a Hilbert space. Due to the spectral decomposition theorem [75, 87], a key role is played by *projection valued measures* (PVM): the set of PVM can be put in a bijective correspondence with the set $\mathcal{A}$ of self adjoint operators of $\mathcal{H}$. Intuitively speaking, a PVM is a map which assigns a projection operator to each interval of the real line. In this sense, projection operators are the building blocks of any other observable in QM. In order to give a precise definition, we need first to specify what we mean by "intervals of the real line". The *borel sets* ($B(\mathbb{R})$) are defined as the family of subsets of $\mathbb{R}$ such that

- it is closed under set theoretical complements,
- it is closed under denumerable unions, and
- it includes all open intervals [75].

Then, a PVM is a map M defined as follows

$$M : B(\mathbb{R}) \to \mathcal{P}(\mathcal{H}), \tag{7a}$$

satisfying

$$M(0) = 0 \tag{7b}$$

$$M(\mathbb{R}) = \mathbf{1} \tag{7c}$$

$$M(\cup_j(B_j)) = \sum_j M(B_j), \tag{7d}$$

for any disjoint denumerable family B_j. Also,

$$M(B^c) = \mathbf{1} - M(B) = (M(B))^{\perp} \tag{7e}$$

Fixing an element $A \in \mathcal{A}$, the intended interpretation of the associated PVM $M_A((a,b)) = P_{(a,b)}$ is "the value of A lies between the interval (a,b)". In this sense, projections represent elementary tests or propositions in QM. As we reviewed in the previous Section, projection operators can

be endowed with a lattice structure, and thus, also elementary tests. This lattice was called "Quantum Logic" by Birkhoff and von Neumann [7]. We will refer to it as the *von Neumann-lattice* ($\mathcal{L}_{v\mathcal{N}}(\mathcal{H})$) [77].
As shown in [7], an analogous treatment can be done for classical systems. In that theory, an observable represented by a real function f in phase space, assigns to each Borel subset E the subset $f^{-1}(E)$ of the phase space Γ. As an example, to the proposition "the value of f lies between the interval (a, b)" it corresponds a subset $f^{-1}(a, b)$ of Γ. Thus, in CM, propositions or elementary tests are associated to subsets of the phase space, and logical connectives "$\vee$", "$\wedge$" and "$\neg$", are represented by *union*, *intersection* and (set theoretical) *complement* of subsets of Γ. The resulting lattice is a Boolean one.

During the thirties, von Neumann and collaborators, continued studying formal developments related to the quantum formalism. One of the results of this investigation was the development of the theory of *rings of operators* (better known as *von Neumann algebras* [77, 65, 66, 88, 67]), as an attempt of generalizing certain algebraic properties of Jordan algebras [58]. The subsequent study of von Neumann algebras showed that they are closely related to lattice theory. Murray and von Neumann provided a classification of factors (von Neumann algebras whose center is formed by the multiples of the identity) using orthomodular lattices in [65, 66, 88, 67]. On the other hand, lattice theory is deeply connected to projective geometry [81], and one of the major discoveries of von Neumann was that of *continuous geometries*, which do not possess "points" (or "atoms") and are related to type II factors. Far from being a mere mathematical curiosity, type II factors found applications in statistical mechanics and type III factors play a key role in the axiomatic approach to Quantum Field Theory (QFT) [76, 77].
The quantum logical approach of Birkhoff and von Neumann was continued by other researchers [45, 69, 60, 85, 86, 35, 2, 3, **?**] (and see [39, 27, 50] for complete expositions). One of the key results of this approach is the *representation theorem* of C. Piron [69]. He showed that any propositional system can be coordinatized in a generalized Hilbert space. A later result by Solèr shows that adding extra assumptions, it can only be a Hilbert space over the fields of the real numbers, complex numbers or quaternions [82].

3.2 Quantal Effects

Projective measures are not the only way in which observable quantities can be described in QM. There exists a more general notion, namely, that of *quantal effect*. And it turns out that this notion can be generalized to arbitrary statistical theories. We will start by reviewing the general case before we consider the quantum one. If the state space of a given probabilistic theory is given by the set Σ, let us denote by X to the set of possible measurement outcomes of an observable quantity. Then, if the system is in a state s, a probability $p(x, s)$ is assigned to any possible outcome $x \in X$. This probability should be well defined in order that our theory be considered as a probabilistic one. In this way, we must have a function

$$\begin{aligned} p : X \times \Sigma &\mapsto [0, 1] \\ (x, s) &\to p(x, s) \end{aligned} \tag{8}$$

To each outcome $x \in X$ and state $s \in \Sigma$, this function assigns a probability $p(x, s)$ of x to occur if the system is in the state s. In this way, a triplet $(\Sigma, p(\cdot\,,\cdot\,), X)$ is assigned for each system of any probabilistic theory [60]. Thinking of s as variable, we obtain the mapping $s \mapsto p(\cdot, s)$ from $\Sigma \to [0, 1]^X$. This implies that all the states of Σ can be identified with maps which generates a canonical vector space. Their closed convex hull forms a new set Ω representing all possible probabilistic mixtures (convex combinations) of states in Σ. Given an arbitrary $\alpha \in \Omega$ and any

outcome $x \in X$, we can define an affine evaluation-functional $f_x : \Omega \to [0,1]$ in a canonical way by $f_x(\alpha) = \alpha(x)$. More generally, we can consider any affine functional $f : \Omega \to [0,1]$ as representing a measurement outcome, and thus use $f(\alpha)$ to represent the probability for that outcome in state α.
Due to the fact that QM is also a probabilistic theory, it follows that it can be included in the general framework described above. In QM, the set of all affine functionals defined as above are called *effects*. The generalization of the notion of PVM (which is based on projections) to an observable based on effects will be called a *positive operator valued measure* (POVM)) [17, 41, 68, 30, 25, 18, 31] and in QM will be represented by a mapping

$$E : B(\mathcal{R}) \to \mathcal{B}(\mathcal{H}). \tag{9a}$$

such that

$$E(\mathcal{R}) = \mathbf{1} \tag{9b}$$

$$E(B) \geq 0, \text{ for any } B \in B(\mathcal{R}) \tag{9c}$$

$$E(\cup_j(B_j)) = \sum_j E(B_j), \text{ for any disjoint familly } B_j \tag{9d}$$

A POVM is thus a measure whose values are non-negative self-adjoint operators on a Hilbert space, and the above definition reduces to the PVM case when these operators are also projections. It is the most general formulation of the description of a measurement in the framework of quantum physics. Positive operators $E(B)$ satisfyng $0 \leq E \leq \mathbf{1}$ are called effects and generate an *effect algebra* [30, 25])and we will denote it by E($\mathcal{H}$). In a similar way as in the general setting, in QM a POVM defines a family of affine functionals on the quantum state space $\mathcal{C}$ of all positive hermitian trace-class operators of trace one (which corresponds to Ω in the general probabilistic setting):

$$E(B) : \mathcal{C} \to [0,1] \tag{10a}$$

$$\rho \mapsto \mathrm{tr}(E\rho) \tag{10b}$$

4 Classical Probabilities

This Section is devoted to *classical probability theory*. But what do we mean with this notion? There exists a vast literature and tendencies disputing the meaning of CP. We will not give a detailed survey of the discussion here, but we discuss two of the most important approaches to CP. These are the one given by A. N. Kolmogorov [57] and the one given by R. T. Cox [23, 22].

4.1 Kolmogorov

Kolmogorov presented his axiomatization of classical probability theory [57] in the 30's as follows. Given an outcome set Ω, let us consider a σ-algebra Σ of subsets of Ω. A probability measure will be a function μ such that

$$\mu : \Sigma \to [0,1] \tag{11a}$$

satisfying

$$\mu(\emptyset) = 0 \tag{11b}$$

$$\mu(A^c) = 1 - \mu(A), \tag{11c}$$

where $(\ldots)^c$ means set-theoretical-complement and for any pairwise disjoint denumerable family $\{A_i\}_{i\in I}$

$$\mu(\bigcup_{i\in I} A_i) = \sum_i \mu(A_i) \tag{11d}$$

Conditions (11) are known as *axioms of Kolmogorov* [57]. The triad (Ω, Σ, μ) is called a *probability space*. Probability spaces obeying Eqs. (11) are usually referred as Kolmogorovian, classical, commutative or boolean probabilities [35].
It is possible to show that if (Ω, Σ, μ) is a Kolmogorovian probability space, the *inclusion-exclusion principle* holds

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \tag{12}$$

or (as expressed in logical terms)

$$\mu(A \vee B) = \mu(A) + \mu(B) - \mu(A \wedge B) \tag{13}$$

As remarked in [74], Eq. (12) was considered as crucial by von Neumann for the interpretation of $\mu(A)$ and $\mu(B)$ as relative frequencies. If $N_{(A\cup B)}$, $N_{(A)}$, $N_{(B)}$, $N_{(A\cap B)}$ are the number of times of each event to occur in a series of N repetitions, then (12) trivially holds.
This principle does no longer hold in QM, a fact linked to the non-boolean QM-character. Thus, the relative-frequencies' interpretation of quantum probabilities becomes problematic. The QM example shows that non-distributive propositional structures play an important role in probability theories *different from* that of Kolmogorov.

4.2 Cox's approach

Since the beginning of probability theory, there has been a school of thought known as Bayesianism, which treated probabilities in a different manner as the one discussed in the previous section. To them, probabilities were not to be regarded as a property of a system, but as a property of our knowledge about it. This position is present as early as in the XIX century in one of the milestones in the development of this theory [59]. In his work, Laplace proposed a way to assign probabilities in situations of ignorance that would eventually be known as "Laplace principle". Later works would try to formalize and give coherence to Bayesian probability [48] [28]. In this section we will center our attention on one of this attempts [23, 22], due to R. T. Cox.
While attaining equivalent results to Kolmogorov, Cox's approach is conceptually very different. In the Kolmogorovian approach probabilities are naturally interpreted (but not necessarily) as relative frequencies in a sample space. On the other hand, in the approach developed by Cox, probabilities are considered as a measure of the degree of belief of an intelligent agent (which may as well be a machine), on the truth of a proposition x if it is known that proposition y is true. In this way, Cox intended to find a set of rules to inferential reasoning that would be coherent with classical logic and would reduce to it in the case of premises with definite truth values.
To do this, he departed from two general axioms and presupposed the calculus of classical propositions which, as is well known, forms a boolean lattice [19]. By doing so, he derived classical probability theory as an inferential calculus on boolean lattices. We will sketch here

the arguments presented in his book [23]. For a more detailed exposition on the deductions, the reader is referred to [22, 23, 46, 89, 52, 53, 55].
The two axioms used by Cox [23] are

- C1 - The probability of an inference on given evidence determines the probability of its contradictory on the same evidence.
- C2 - The probability on a given evidence that both of two inferences are true is determined by their separate probabilities, one on the given evidence, the other on this evidence with the additional assumption that the first inference is true.

A real valued function φ representing the degree to which a proposition h (usually called *hypothesis*) implies another proposition a is postulated. Thus, $\varphi(a|h)$ will represent the degree of belief of an intelligent agent of how likely it is that a is true given that he knows that the hypothesis h is true.
Then, requiring the function φ to be coherent with the properties of the calculus of classical propositions, he derives the rules for manipulating probabilities. Using axiom C2, the associativity of the conjunction $(a \wedge (b \wedge c) = (a \wedge b) \wedge c))$, and defining the function $F[\varphi(a|h), \varphi(a|h)] \equiv \varphi(a \wedge b|h) : \mathbb{R}^2 \to \mathbb{R}$; he arrives at a functional equation for $F(x, y)$:

$$F[x, F(y, z)] = F[F(x, y), z] \tag{14}$$

Which, after a rescaling and a proper definition of the probability $P(a|h)$ in terms of $\varphi(a|h)$, leads to the well known *product rule* of probability theory:

$$P(a \wedge b|h) = CP(a|h \wedge b)P(b|h) \tag{15}$$

The definition of $P(a|h)$ in terms of $\varphi(a|h)$ is omitted, as one ultimately ends up using only the function $P(a|h)$ and never $\varphi(a|h)$. In an analogous manner, using axiom C1, the law of double negation $(\neg\neg a = a)$, Morgan's law for disjunction $(\neg(a \vee b) = \neg a \wedge \neg b)$ and defining the function $f[P(a|h)] \equiv P(\neg a|h) : \mathbb{R} \to \mathbb{R}$, he arrives at the following functional equation for $P(a|h)$

$$P(a|h)^r + P(\neg a|h)^r = 1 \tag{16}$$

With r an arbitrary constant. Although in principle different values of r would give rise to different rules for the computation of the probability of the negation of a proposition, as taking different values of r account for a rescaling of $P(a|h)$, one could as well call $P'(a|h) \equiv P(a|h)^r$ *probability* and work with this function instead of $P(a|h)$. For simplicity, Cox decides to take $r = 1$ and continue using $P(a|h)$.
Using equations (14), (16), the law of double negation and Morgan's law for conjunction $(\neg(a \wedge b) = \neg a \vee \neg b)$, he arrives to the *sum rule* of probability theory:

$$P(a \vee b|h) = P(a|h) + P(b|h) - P(a \wedge b|h) \tag{17}$$

As it turns out, $P(a|h)$ –if suitably normalized– satisfies all the properties of a Kolmogorovian probability (Eqs. (11)).
Because of the importance of Cox's theorem to the foundations of probability, it has been target of a thorough scrutiny by many authors. Some have pointed out inconsistencies behind the implicit assumptions made during its derivations, most notably the assumptions behind the validity of equation 14. Since then, there have been different proposals to save Cox's approach by proving it through less restrictive axioms. In [40], a discussion of the status of Cox proposal

is presented as well as a counterexample to it. For a review on the subject, it is recommended to consult [89].
Once the general properties of the function $P(a|h)$ are established, the next problem is to find a way to determine prior probabilities (probabilities conditional only to the hypothesis h). Although formally one could assign prior probabilities in any way coherent with the normalization used [73], in practical situations one is compelled to assign them in a way that they reflect the information contained in the hypothesis h. A possible way to do this is by using the MaxEnt principle[46][47], which we will review shortly in the next Section. Other ways of assigning prior probabilities include the Laplace principle [48] and coherence with symmetry transformations [49]. Nevertheless, the existance of a general algorithm for assigning prior probabilities is still an open question.

4.3 MaxEnt Principle

This principle asserts that the assignment of the prior probabilities from a hypothesis h should be done by maximizing the uncertainty associated with its distribution while respecting the constrains imposed over them by h. Although this may sound paradoxical, by maximizing the uncertainty of the prior probabilities one avoids to assume more information than the one strictly contained in h.
Taking Shannon's information measure $S[P] = -\sum_i P(a_i|h) log[P(a_i|h)]$ as the measure of the uncertainty associated with the distribution P, MaxEnt principle can be restated as: the prior probabilities corresponding to the hypothesis h are given by the distribution that maximizes $S[P]$ subject to the constraints imposed by h on P. The simplest example is given by the hypothesis h that imposes no constraints on P, in which case P results to be the uniform distribution and the MaxEnt principle reduces to Laplace's. Different kinds of constraints result on different prior probability distributions (PPD) [47]. In [72] a table of some of the distributions obtained this way is presented. It must be pointed out that, although given a set of constraints its corresponding PPD can readily be computed, there is no general way of translating a hypothesis h into equivalent constraints.
By means of the MaxEnt principle, classical and quantum equilibrium statistical mechanics can be formulated on the basis of information theory [46]. Assuming that your prior knowledge about the system is given by the values of n expectation values of physical quantities R_j, i.e., $\langle R_1\rangle, \ldots, \langle R_n\rangle$, then the most unbiased probability distribution $\rho(x)$ is uniquely fixed by extremizing Shannon's logarithmic entropy S subject to the n constraint

$$\langle R_i\rangle = r_i; \text{ for all } i. \tag{18}$$

In order to solve this problem, n Lagrange multipliers λ_i must be introduced.

In the process of employing the MaxEnt procedure one discovers that the information quantifier S can be identified with the equilibrium entropy of thermodynamics if our prior knowledge $\langle R_1\rangle, \ldots, \langle R_n\rangle$ refers to extensive quantities [46]. $S(maximal)$, once determined, yields complete thermodynamical information with respect to the system of interest [46]. The MaxEnt probability distribution function (PDF), associated to Boltzmann-Gibbs-Shannon's logarithmic entropy S, is given by [46]

$$\rho_{max} = \exp\left[(-\lambda_0 \mathbf{1} - \lambda_1 R_1 - \cdots - \lambda_n R_n)\right], \tag{19}$$

where the λ's are Lagrange multipliers guaranteeing that

$$r_i = -\frac{\partial}{\partial \lambda_i} \ln Z, \tag{20}$$

while the partition function reads

$$Z(\lambda_1 \cdots \lambda_n) = \mathrm{tr}[\exp^{-\lambda_1 R_1 - \cdots - \lambda_n R_n}], \tag{21}$$

and the normalization condition

$$\lambda_0 = \ln Z. \tag{22}$$

In a quantum setting, the R's are operators on a Hilbert space $\mathcal{H}$ while ρ is a density matrix (operator).

5 Convex Operational Models

As seen in Section 3.2, an arbitrary statistical theory can be endowed with a convex set of states Ω and a set of generalized observables (effects) in a natural way. Generalized observables will be reasonably represented by affine functionals and we will call $A(\Omega)$ to the space of all affine functionals. This key observation leads to a general approach to statistical theories based on the study of the geometrical properties of convex sets. This is the starting point of the COM approach. In this Section we concentrate on elementary notions of COM's, and we refer the reader to [11] for en excellent presentation of the subject (and we follow them in this Section).
We saw that a probability $a(\omega) \in [0,1]$ is well defined for any state $\omega \in \Omega$ and an observable a. In the COM approach, it is usually assumed that there exists a unitary observable u such that $u(\omega) = 1$ for all $\omega \in \Omega$. Thus, in analogy with the quantum case, the set of all effects will be encountered in the interval $[0, u]$ (the order in the general case is the canonical one in the sapce of affine functionals). A (discrete) measurement will be represented by a set of effects $\{a_i\}$ such that $\sum_i a_i = u$. Ω can be naturally embedded in the dual space $A(\Omega)^*$ using the map

$$\begin{aligned} \omega &\mapsto \hat{\omega} \\ \hat{\omega}(a) &:= a(\omega) \end{aligned} \tag{23}$$

Let $V(\Omega)$ be the linear span of Ω in $A(\Omega)^*$. Then, it is reasonable to consider Ω finite dimensional if and only if $V(\Omega)$ is finite dimensional. For the sake of simplicity, we will restrict ourselves to this case (and to compact spaces). As is well known, this implies that Ω can be expressed as the convex hull of its extreme points. the extreme points will represent *pure states* (in the QM case, pure quantum states are indeed the extreme points of $\mathcal{C}$). It can be shown that for finite dimension d, a system will be classical if and only if it is a simplex[1]. It is a well known fact that in a simplex a point may be expressed as a unique convex combination of its extreme points. This characteristic feature of classical theories no longer holds in quantum models. Indeed, in the case of QM, there are infinite ways in which one can express a mixed as a convex combination of pure states (for a graphical representation, think about the maximally mixed state in the Bloch sphere).
Interestingly enough, there is also a connection between the faces of the convex set of states of a given model and its lattice of properties (in the quantum-logical sense), providing an unexpected connection between geometry, lattice theory and statistical theories. Faces of a convex set are defined as subsets which are stable under mixing and purification. This is means that F is a face if each time that

$$x = \lambda x_1 + (1-\lambda) x_2, \quad 0 \leq \lambda \leq 1, \tag{24}$$

[1] A simplex is the convex hull of $d+1$ linearly independent pure states.

then $x \in F$ if and only if $x_1 \in F$ and $x_2 \in F$ [20]. The set of faces of a convex set form a lattice in a canonical way and it can be shown that the lattice of faces of a classical model is a boolean one. On the other hand, in QM, the lattice of faces of the convex set of states $\mathcal{C}$ (defined as the set of positive trace class hermitian operators of trace one), is isomorphic to the von Neumann lattice of closed subspaces $\mathcal{P}(\mathcal{H})$ [20, 8].

Let us turn now to compound systems. Given a compound system, its components will have state spaces Ω_A and Ω_B. Let us denote the joint state space by Ω_{AB} . It is reasonable to identify Ω_{AB} with the linear span of $(V(\Omega_A) \otimes V(\Omega_B))$ [11]. Then, a maximal tensor product state space $\Omega_A \otimes_{max} \Omega_B$ can be defined as the one which contains all bilinear functionals $\varphi : A(\Omega_A) \times A(\Omega_B) \longrightarrow \mathbb{R}$ such that $\varphi(a, b) \geq 0$ for all effects a and b and $\varphi(u_A, u_B) = 1$. The maximal tensor product state space has the property of being the biggest set of states in $(A(\Omega_A) \otimes A(\Omega_B))^*$ which assigns probabilities to all product- measurements. The minimal tensor product state space $\Omega_A \otimes_{min} \Omega_B$ is simply defined by the convex hull of all product states. A product state will then be a state of the form $\omega_A \otimes \omega_B$ such that

$$\omega_A \otimes \omega_B(a, b) = \omega_A(a)\omega_B(b), \tag{25}$$

for all pairs $(a, b) \in A(\Omega_A) \times A(\Omega_B)$. Given a particular compound system of a general statistical theory, its set of states Ω_{AB} —we will call it $\Omega_A \otimes \Omega_B$ from now on— will satisfy

$$\Omega_A \otimes_{min} \Omega_B \subseteq \Omega_A \otimes \Omega_B \subseteq \Omega_A \otimes_{max} \Omega_B \tag{26}$$

As expected, for classical compound systems (because of the absence of entangled states), we will have $\Omega_A \otimes_{min} \Omega_B = \Omega_A \otimes_{max} \Omega_B$. For the quantum case we have the strict inclusions $\Omega_A \otimes_{min} \Omega_B \subseteq \Omega_A \otimes \Omega_B \subseteq \Omega_A \otimes_{max} \Omega_B$. The general definition of a separable state in an arbitrary COM is made in analogy with that of [91], i.e., as one which can be written as a convex combination of product states [12, 71]:

Definition 5.1. *A state $\omega \in \Omega_A \otimes \Omega_B$ will be called* separable *if there exist $p_i \in \mathbb{R}_{\geq 0}$, $\omega_A^i \in \Omega_A$ and $\omega_B^i \in \Omega_B$ such that*

$$\omega = \sum_i p_i \omega_A^i \otimes \omega_B^i, \quad \sum_i p_i = 1 \tag{27}$$

If $\omega \in \Omega_A \otimes \Omega_B$ is not separable, then, it will be reasonably called *entangled* [20, 83, 84]. As expected, entangled states exist only if $\Omega_A \otimes \Omega_B$ is strictly greater than $\Omega_A \otimes_{min} \Omega_B$.

The COM approach already shows that, given an arbitrary statistical theory, there is a generalized notion of probabilities of measurement outcomes. These probabilities are encoded on the states in Ω. And we have seen that there are many differences between classical state spaces and non-classical ones: this is expressed in the geometrical properties of their convex state spaces and in the correlations appearing when compound systems are considered. Indeed, QM and classical probability theories are just particular COMs among a vast family of possibilities. In the next Section, we will study with more detail the axiomatization of the probabilities appearing in non-classical models.

6 Quantum Probabilities

In this Section we discuss QP. We start first by reviewing the usual approach, in which Kolmogorov's axioms are extended to non-Boolean lattices (or algebras) [76]. Next, we review a relatively recent approach to the probabilities appearing in QM which uses distributive lattices [37, 36, 56].

6.1 The Orthodox Approach

As we have seen in Section 3.1 , elementary tests in QM are represented by closed subspaces of Hilbert space. These subspaces form an orthomodular atomic lattice $\mathcal{P}(\mathcal{H})$. In order to assign probabilities to these elementary tests or processes, many texts proceed by postulating axioms which are similar to those of Kolmogorov [77, 87, 8]. The Boolean Σ-algebra appearing in Kolmogorov's axioms (Eqn. (11)) is replaced by $\mathcal{P}(\mathcal{H})$:

$$s : \mathcal{P}(\mathcal{H}) \to [0; 1] \tag{28a}$$

such that:

$$s(\mathbf{0}) = 0 \ (\mathbf{0} \text{ is the null subspace}). \tag{28b}$$

$$s(P^{\perp}) = 1 - s(P), \tag{28c}$$

and, for a denumerable and pairwise orthogonal family of projections

$$P_j, \ s(\sum_j P_j) = \sum_j s(P_j). \tag{28d}$$

In this way, a real number between 0 and 1 is assigned to any elementary test. Despite of the similarity with Kolmogorov's axioms, the probabilities defined above are very different, due to the non-Boolean character of the lattice involved. This feature led to many authors to assert that there are no quantum probabilities (just because they are not Kolmogorovian). Gleason's theorem [32, 21] asserts that if the dimension of $\mathcal{H} \geq 3$, any measure s satisfying (28) can be put in correspondence with a trace class operator (of trace one) ρ_s:

$$s(P) := \mathrm{tr}(\rho_s P) \tag{29}$$

On the other hand, using equation (29) any trace class operator of trace one defines a measure as in (28). In this way, equations (28) define the usual probabilities of QM.
We have mentioned above that (11) and (28) are not equivalent probability theories. For example, Eq. (12) is no longer valid in QM. Indeed, for suitably chosen quantum events A and B we have

$$s(A) + s(B) \leq s(A \vee B) \tag{30}$$

The probability theory defined by (28) may be considered as a generalization of classical probability theory in the following sense: while in an arbitrary statistical theory a state will be a normalized measure over a suitable C^*-algebra, the classical case is recovered when the algebra is *commutative* [35, 76]. Difficulties which appear when one tries to define a quantum conditional probability. For a complete discussion about these matters and a comparison between classical and quantum probabilities, see [35] and [76].

6.2 Cox's Method Applied To Physics

A novel derivation of Feynman's rules for quantum mechanics was presented in refs. [37], [36], and [56]. An experimental *logic of processes* for quantum systems is presented, and this is done in such a way that the resulting lattice is a distributive one. This is a major difference with the approach described in Section 6.1, because the lattice of projections in a Hilbert space is non-distributive.

The logic of processes is constructed as follows. Given a sequence of measurements M_1,...,M_n on a quantum system, yielding results m_1, m_2, ..., m_n, a particular process is represented as a *measuring sequence* $A = [m_1, m_2, \ldots, m_n]$.
Next, conditional (logical) propositions $[m_2, \ldots, m_n|m_1]$ are introduced. Using them, a probability is naturally associated to a sequence A with the formula

$$P(A) = Pr(m_n, \ldots, m_2|m_1) \tag{31}$$

representing the probability of obtaining outcomes m_2, ..., m_n conditional upon obtaining m_1. Let us see how this works with a concrete example in which the m_i's has two possible values, 0 and 1. Then $A_1 = [0, 1, 1]$ and $A_2 = [0, 0, 1, 1]$ represent measuring sequences of three and four measurements respectively. P(A)=Pr(1,1—0), represents the probability of obtaining outcomes $m_2 = 1$ and $m_3 = 1$ conditional upon obtaining $m_1 = 0$.
Measurements can be coarse grained as follows. Suppose that we want to coarse grain M_2. Then, we can unite the two outcomes 0 and 1 in a joint outcome $(0, 1)$. Then, a new measurement $\widetilde{M_2}$ is created. Thus, a possible sequence obtained by the replacement of M_2 by $\widetilde{M_2}$ could be $[1, (1, 2), 1]$. Analogous constructions can be done for other measurements. In this way an operation can be defined for sequences:

$$[m_1, \ldots, (m_i, m_i'), \ldots, m_n] := [m_1, \ldots, m_i, \ldots, m_n] \vee [m_1, \ldots, m_i', \ldots, m_n] \tag{32}$$

And another operation can be defined reflecting the fact that sequences can be compounded as follows

$$[m_1, \ldots, m_j, \ldots, m_n] := [m_1, \ldots, m_j] \cdot [m_j, \ldots, m_n] \tag{33}$$

With these operations at hand, it is easy to show that if A, B and C are measuring sequences, then

$$A \vee B = B \vee A \tag{34a}$$

$$(A \vee B) \vee C = A \vee (B \vee C) \tag{34b}$$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \tag{34c}$$

$$(A \vee B) \cdot C = (A \cdot C) \vee (B \cdot C) \tag{34d}$$

$$C \cdot (A \vee B) = (C \cdot A) \vee (C \cdot B), \tag{34e}$$

Equations (34) show explicitly that "$\vee$" is commutative and associative, "$\cdot$" is associative, and that there is right- and left-distributivity of "$\cdot$" over "$\vee$".
Equations (34) define the algebraic "symmetries" of the experimental logic of processes. As in the approach of Cox to classical probability, these symmetries are used to derive Feynman's rules [37]. But at this step, a crucial assumption is made: each measuring sequence will be represented by a pair of real numbers [37]. This assumption is justified in [37] by appealing to Bohr's complementarity principle.
If measuring sequences A, B, etc. induce pairs of real numbers $\mathbf{a}$, $\mathbf{b}$, etc., then, due to equations (34), the associated real numbers should satisfy

$$\mathbf{a} \vee \mathbf{b} = \mathbf{b} \vee \mathbf{a} \tag{35a}$$

$$(\mathbf{a} \vee \mathbf{b}) \vee \mathbf{c} = \mathbf{a} \vee (\mathbf{b} \vee \mathbf{c}) \tag{35b}$$

$$(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c}) \tag{35c}$$

$$(\mathbf{a} \vee \mathbf{b}) \cdot \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) \vee (\mathbf{b} \cdot \mathbf{c}) \tag{35d}$$

$$\mathbf{c} \cdot (\mathbf{a} \vee \mathbf{b}) = (\mathbf{c} \cdot \mathbf{a}) \vee (\mathbf{c} \cdot \mathbf{b}) \tag{35e}$$

The reader can easily verify that equations (35) are satisfied by the field of complex numbers (provided that the operations are interpreted as sum and product of complex numbers). How can we assure that complex numbers are the only field which satisfies equations (35)? In order to single out complex numbers among other possible fields, additional assumptions must be added, namely, pair symmetry, additivity and symmetric bias (see [37, 36, 56] for details). Once these conditions are assumed, the path is clear to derive Feynman's rules by applying a deduction similar to that of Cox, to the experimental logic defined by equations (34).

7 Generalization of Cox's method

As we have seen in previous Sections, there are two versions of CP, namely, the approach of R. T. Cox [23, 22] and the one of A. N. Kolmogorov [57]. The Kolmogorovian approach can be generalized in order to include non-Boolean models. In what follows, we will see that Cox's method can also be generalized to non-distributive lattices, and thus the non-commutative character of QP can be captured in this framework [44, 42].

7.1 Generalized probability calculus using Cox's method

As we have seen in Section 4, Cox studies the functions defined over a distributive lattices and derives classical probabilities. In [44] it is shown that if the lattice is assumed to be non-distributive, the properties of QP described in Section 6.1 can be derived by applying a variant of the Cox's method as follows (see [44]). Suppose that propositions of our system are represented by the lattice of elementary tests of QM, i.e., the lattice $\mathcal{L}_{v\mathcal{N}}(\mathcal{H})$. This lattice isomorphic to the lattice of projections $\mathcal{P}(\mathcal{H})$ of the Hilbert space $\mathcal{H}$. The goal is to show that the "degree of implication" measure $s(\cdots)$ demanded by Cox's method satisfies Eqs. (28). This means that we are looking for a function to the real numbers s, such that it is non-negative and $s(P) \leq s(Q)$ whenever $P \leq Q$.
The operation "$\vee$" in $\mathcal{P}(\mathcal{H})$ is associative. Then, if P and Q are orthogonal projections, the relationship between $s(P)$, $s(Q)$, and $s(P \vee Q)$ must be of the form

$$s(P \vee Q) = F(s(P), s(Q)), \tag{36}$$

with F a function to be determined. If a third proposition"$\vee$" and following a similar procedure to that of Cox, the following functional equation is found

$$F(F(s(P), s(Q)), s(R)) = F(s(P), F(s(Q), s(R))). \tag{37}$$

The above equation can be solved up to rescaling [52, 53, 55, 1], and we find

$$s(P \vee Q) = s(P) + s(Q). \tag{38}$$

whenever $P \perp Q$. It can be shown that for any finite family of orthogonal projections P_j, $1 \leq j \leq n$ [44]:

$$s(\bigvee_{j=1}^{\infty} P_j) = \sum_{j=1}^{\infty} s(P_j), \tag{39}$$

and we recover condition (28d) of the axioms of quantum probability. By exploiting the properties of the orthogonal complement acting on subspaces, it can also be shown [44] that

$$s(P^{\perp}) = 1 - s(P), \tag{40}$$

which is nothing but condition (28c). On the other hand, as $\mathbf{0} = \mathbf{0} \vee \mathbf{0}$ and $\mathbf{0} \perp \mathbf{0}$, then $s(\mathbf{0}) = s(\mathbf{0}) + s(\mathbf{0})$, and thus, $s(\mathbf{0}) = 0$, which is condition (28b). In this way, the Cox's method applied to the non-distributive lattice $\mathcal{L}_{v\mathcal{N}}(\mathcal{H})$ (or $\mathcal{P}(\mathcal{H})$) yields the same probability theory as the one provided by equations (28) for the quantum case.
What happens if the Cox's method is applied to an arbitrary atomic orthomodular lattice $\mathcal{L}$? Now we must define a function $s : \mathcal{L} \longrightarrow \mathbb{R}$, such that it is always non-negative $s(a) \geq 0 \ \forall a \in \mathcal{L}$ and is also order preserving $a \leq b \longrightarrow s(a) \leq s(b)$. In [44] it is shown that under these rather general assumptions, in *any* orthomodular lattice and for any orthogonal denumerable family $\{a_i\}_{i \in \mathbb{N}}$, s must satisfy (up to rescaling)

$$s(\bigvee \{a_i\}_{i \in \mathbb{N}}) = \sum_{i=1}^{\infty} s(a_i) \tag{41a}$$

$$s(\neg a) = 1 - s(a) \tag{41b}$$

$$s(\mathbf{0}) = 0. \tag{41c}$$

In this way, a generalized probability theory is derived (as in (28)). Equations (41) define non-classical (non-Kolmogorovian) probability measures, due to the fact that in any non-distributive orthomodular lattice, there always exist elements a and b such that

$$(a \wedge b) \vee (a \wedge \neg b) < a, \tag{42}$$

But in any classical probability theory, $s(a \wedge \neg b) + s(a \wedge b) = s(a)$ is always satisfied.

8 Conclusions

We have presented a new approach to probabilities appearing in QM. While there exist two alternative formalisms to CP (the Kolmogorovian and the one due to R. T. Cox), we have also shown that these two approaches can be extended to the non-commutative case. In this way, we find that CP are just a particular case of a more general mathematical framework in which, the lattice is distributive. QP is also a particular case of a vast family of theories for which the propositional lattice is non-distributive. Thus, we have a precise formal expression of the notion of QP. Of course, these formal frameworks does not exhaust the philosophical debate around the existence or not of a well defined notion of QP; notwithstanding, the extension of the Cox's method to the non-distributive case and the possibility of including a description of the probabilities in QM in it, constitutes a precise step towards understanding the notion of QP offering a new point of view of this notion.

Acknowledgements This work was partially supported by the grants PIP N° 6461/05 amd 1177 (CONICET). Also by the projects FIS2008-00781/FIS (MICINN) - FEDER (EU) (Spain, EU).

References

[1] Aczél, J., *Lectures on Functional Equations and Their Applications*, Academic Press, New York, (1966).

[2] D. Aerts and I. Daubechies, Lett. Math. Phys. **3** (1979) 11-17.

[3] D. Aerts and I. Daubechies, Lett. Math. Phys. **3** (1979) 19-27.

[4] D. Aerts, J. Math. Phys. **24** (1983) 2441.

[5] D. Aerts, Rep. Math. Phys **20** (1984) 421-428.

[6] D. Aerts, J. Math. Phys. **25** (1984) 1434-1441.

[7] G. Birkhoff and J. von Neumann, Annals Math.**37** (1936) 823-843.

[8] E. G. Beltrametti and G. Cassinelli, *The Logic of Quantum Mechanics* (Addison-Wesley, Reading, 1981).

[9] H. Barnum and A. Wilce, Electronic Notes in Theoretical Computer Science Volume **270**, Issue 1, Pages 3-15,(2011).

[10] H. Barnum, R. Duncan and A Wilce, arXiv:1004.2920v1 [quant-ph] (2010).

[11] H. Barnum, J. Barrett, M. Leifer, and A.Wilce, *Phys. Rev. Lett.* **99**, 240501 (2007).

[12] H. Barnum, O. C. O. Dahlsten, M. Leifer, and B. Toner, in Information Theory Workshop, 2008, pp. 386-390, (2008).

[13] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce and R. Wilke, 2010 New J. Phys. **12**, 033024 (2010), and references therein.

[14] E. Beltrametti, S. Bugajski and V. Varadarajan, J. Math. Phys. **41** (2000)

[15] J. Barrett, Phys. Rev. A 75, 032304 (2007).

[16] H. Barnum, J. Barret, M. Leifer and A. Wilce, *Proceedings of Symposia in Applied Mathematics*, **71**, (2012) 25-47.

[17] P. Busch, J. Kiukas and P. Lahti, arXiv:0905.3222v1 [quant-ph] (2009).

[18] P. Busch, P. Lahti, P. Mittlestaedt, *The Quantum Theory of Measurement* (Springer-Verlag, Berlin, 1991).

[19] Boole, G. *An Investigation of the Laws of Thought*, Macmillan: London, UK, (1854).

[20] I. Bengtsson, K. Zyczkowski, Geometry of Quantum States: An Introduction to Quantum Entanglement, Cambridge Univ. Press, Cambridge, 2006.

[21] D. Buhagiar, E. Chetcuti and A. Dvurečenskij, Found. Phys. **39**, 550-558 (2009).

[22] Cox, R.T. Probability, frequency, and reasonable expectation. Am. J. Phys. 14, 1-13 (1946).

[23] Cox, R.T. The Algebra of Probable Inference; The Johns Hopkins Press: Baltimore, MD, USA, (1961).

[24] A. Caticha, *Phys. Rev. A* **57** (1998) 1572-1582.

[25] G. Cattaneo and S. Gudder, Found. Phys. **29**, 1607-1637,(1999)

[26] E. Davies and J. Lewis, Commun. Math. Phys. **17**, (1970) 239-260.

[27] M. L. Dalla Chiara, R. Giuntini, and R. Greechie, *Reasoning in Quantum Theory* (Kluwer Acad. Pub., Dordrecht, 2004).

[28] B. de Finetti, *Teoria Delle Probabilita*, Giulio Einaudi editore, (1970).

[29] A Dvurečenskij and S. Pulmannová, *New Trends in Quantum Structures* (Kluwer Acad. Pub., Dordrecht, 2000.

[30] D. Foulis and S. Gudder, Found. Phys. **31**, 1515-1544, (2001).

[31] D. Foulis, M. K. Bennett, Found. Phys. **24** (1994) 1331-1352.

[32] A. Gleason, J. Math. Mech. **6**, 885-893 (1957).

[33] J. R. Greechie, in *Current Issues in Quantum Logic*, E. Beltrameti and B. van Fraassen, eds. (Plenum, New York, 1981) pp. 375-380.

[34] R. Giuntini, *Quantum Logic and Hidden Variables* (BI Wissenschaftsverlag, Mannheim, 1991).

[35] S. P. Gudder, *Stochastic Methods in Quantum Mechanics* North Holland, New York - Oxford (1979).

[36] Goyal, P., Knuth, K.H. and Skilling, J., *Phys. Rev. A* **81** 022109 (2010).

[37] P. Goyal and K. Knuth, Symmetry, **3**, 171-206 (2011) doi:10.3390/sym3020171.

[38] F. Holik and A. Plastino, in *Contemporary Research in Quantum Systems*, Z. Ezziane ed., NOVA Publishers, New York, (2014).

[39] *Handbook Of Quantum Logic And Quantum Structures* (Quantum Logic), Edited by K. Engesser, D. M. Gabbay and D. Lehmann, North-Holland (2009).

[40] J. Y. Halpern, *Journal of Articial Intelligence Research*, **10**, 67-85, (1999).

[41] T. Heinonen, *Imprecise Measurements In Quantum Mechanics*, PhD. Thesis (2005).

[42] F. Holik, "Logic, Geometry and Probability Theory", *SOP Transactions On Thoeretical Physics*, *in press*, (2014).

[43] F. Holik, C. Massri and A. Plastino, *SOP Transactions on Theoretical Physics*, in press, (2014).

[44] F. Holik, A. Plastino and M. Sáenz, *Annals Of Physics*, Volume **340**, Issue **1**, 293-310, (2014).

[45] J. M. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Cambridge, 1968).

[46] E. T. Jaynes, Phys. Rev. Vol. 106, Number 4 (1957); Phys. Rev. Vol. 108, Number 2 (1957).

[47] E. T. Jaynes, *Probability: the logic of science*, Cambridge University Press, Cambridge (2003).

[48] H. Jeffreys, *Theory of Probability*, Clarendon Press, Oxford, (1939).

[49] E. T. Jaynes, *Foundations of Physics* **3**, 477-493, (1973).

[50] G. Kalmbach, *Orthomodular Lattices* (Academic Press, San Diego, 1983).

[51] G. Kalmbach, *Measures and Hilbert Lattices* (World Scientific, Singapore, 1986).

[52] Knuth, K.H. Deriving laws from ordering relations. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Proceedings of 23rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering; Erickson, G.J., Zhai, Y., Eds.; American Institute of Physics: New York, NY, USA, pp. 204-235 (2004).

[53] Knuth, K.H. Measuring on lattices. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Proceedings of 23rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering; Goggans, P., Chan, C.Y., Eds.; American Institute of Physics: New York, NY, USA,; Volume 707, pp. 132-144 (2004).

[54] Knuth, K.H. Valuations on lattices and their application to information theory. In Proceedings of the 2006 IEEE World Congress on Computational Intelligence, Vancouver, Canada, July (2006).

[55] Knuth, K.H. Lattice duality: The origin of probability and entropy. Neurocomputing, 67C, 245-274 (2005).

[56] Knuth, K.H. and Skilling, arXiv:1008.4831v1 (2012).

[57] Kolmogorov, A.N. Foundations of Probability Theory; Julius Springer: Berlin, Germany, (1933).

[58] J. Lacki, *Arch. Hist. Exact Sci.* **54** (2000) 279-318.

[59] P. S. Laplace, *Essai philosophique sur les probabilités*, (1814).

[60] G. Mackey *Mathematical foundations of quantum mechanics* New York: W. A. Benjamin (1963).

[61] G. W. Mackey, Amer. Math. Monthly, Supplement **64** (1957) 45-57.

[62] B. Mielnik, Commun. Math. Phys. **9** (1968) 55-80.

[63] B. Mielnik, Commun. Math. Phys. **15** (1969) 1-46.

[64] B. Mielnik, Commun. Math. Phys. **37** (1974) 221-256.

[65] F. J. Murray and J. von Neumann, "On rings of operators", *Ann. Of Math.*, **37**, (1936) 116-229.

[66] F. J. Murray and J. von Neumann, "On rings of operators II", *Trans. Amer. Math. Soc.*, **41** (1937), 208-248.

[67] F. J. Murray and J. von Neumann, "On rings of operators IV", Ann. Of Math., **44**, (1943) 716-808.

[68] Z. Ma, arXiv:0811.2454v1 [quant-ph] (2008).

[69] C. Piron, *Foundations of Quantum Physics* (Addison-Wesley, Cambridge, 1976).

[70] P. Pták and S. Pulmannova, *Orthomodular Structures as Quantum Logics* (Kluwer Academic Publishers, Dordrecht, 1991).

[71] Physical Review Letters, vol. 108, Issue 12, (2012) 120502.

[72] S. Y. Park and A. K. Bera, *Journal of Econometrics*, **150**, 219-230, (2009).

[73] G. C. Rota, "On the combinatorics of the Euler characteristic", *Studies in Pure Mathematics*, Academic Press, London, 221-233 (1971).

[74] M. Redei "The Birkhoff-von Neumann Concept of quantum Logic", in *Handbook of Quantum Logic and Quantum Structures*, K. Engesser, D. M. Gabbay and D. Lehmann, eds., Elsevier (2009).

[75] M. Reed and B. Simon, *Methods of modern mathematical physics* I: Functional analysis, Academic Press, New York-San Francisco-London (1972).

[76] M. Rédei and S. Summers, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* Volume **38**, Issue **2**, (2007) 390-417.

[77] M. Rédei, *Quantum Logic in Algebraic Approach* (Kluwer Academic Publishers, Dordrecht, 1998).

[78] M. Srinivas, J. Math. Phys. **16**, (1975) 1672.

[79] K. Svozil, *Quantum Logic* (Springer-Verlag, Singapore, 1998).

[80] A. J. Short and S. Wehner, New J. Phys. 12, 033023 (2010).

[81] I. Stubbe and B.V. Steirteghem *Propositional systems, Hilbert lattices and generalized Hilbert spaces*, in Handbook Of Quantum Logic Quantum Structures: Quantum Structures, K. Engesser, D.M. Gabbay and D. Lehmann, eds., (2007).

[82] M. Solèr, *Communications in Algebra* **23**, 219-243.

[83] E. Schrödinger, Proc. Cambridge Philos. Soc. 31, 555 (1935).

[84] E. Schrödinger, Proc. Cambridge Philos. Soc. 32, 446 (1936).

[85] V. Varadarajan, *Geometry of Quantum Theory I* (van Nostrand, Princeton, 1968).

[86] V. Varadarajan, *Geometry of Quantum Theory II* (van Nostrand, Princeton, 1970).

[87] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, 12th. edition, Princeton, 1996).

[88] J. von Neumann, "On rings of operators III", *Ann. Math.*, **41**, (1940), 94-161.

[89] K. S. Van Horn, *Internat. J. Approx. Reason.*, **34**, 3-24, (2003).

[90] A. Wilce, *Quantum Logic and Probability Theory*, The Stanford Encyclopedia of Philosophy (Spring 2009 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2009/entries/qt-quantlog/. Archive edition: Spring 2009.

[91] R. Werner, Phys. Rev. A **40**, (1989) 4277-4281