Annals of Physics 340 (2014) 293–310

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: www.elsevier.com/locate/aop

A discussion on the origin of quantum probabilities

Federico Holik [a,b,*], Manuel Sáenz [b], Angel Plastino [c]

[a] Universidad Nacional de La Plata, Instituto de Física (IFLP-CCT-CONICET), C.C. 727, 1900 La Plata, Argentina

[b] Departamento de Matemática - Ciclo Básico Común, Universidad de Buenos Aires - Pabellón III, Ciudad Universitaria, Buenos Aires, Argentina

[c] Universitat de les Illes Balears and IFISC-CSIC, 07122 Palma de Mallorca, Spain

HIGHLIGHTS

- Several recent works use a derivation similar to that of R.T. Cox to obtain quantum probabilities.
- We apply Cox's method to the lattice of subspaces of the Hilbert space.
- We obtain a derivation of quantum probabilities which includes mixed states.
- The method presented in this work is susceptible to generalization.
- It includes quantum mechanics and classical mechanics as particular cases.

ARTICLE INFO

Article history:
Received 26 June 2013
Accepted 9 November 2013
Available online 12 November 2013

Keywords:
Quantum probability
Lattice theory
Information theory

ABSTRACT

We study the origin of quantum probabilities as arising from non-Boolean propositional-operational structures. We apply the method developed by Cox to non distributive lattices and develop an alternative formulation of non-Kolmogorovian probability measures for quantum mechanics. By generalizing the method presented in previous works, we outline a general framework for the deduction of probabilities in general propositional structures represented by lattices (including the non-distributive case).

© 2013 Elsevier Inc. All rights reserved.

* Corresponding author at: Universidad Nacional de La Plata, Instituto de Física (IFLP-CCT-CONICET), C.C. 727, 1900 La Plata, Argentina.
E-mail address: olentiev2@gmail.com (F. Holik).

0003-4916/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.aop.2013.11.005

1. Introduction

Quantum probabilities[1] posed an intriguing question from the very beginning of quantum theory. It was rapidly realized that probability amplitudes of quantum process obeyed rules of a non classical nature, as for example, the sum rule of probability amplitudes giving rise to interference terms or the nonexistence of joint distributions for noncommuting observables. In 1936 von Neumann wrote the first work ever to introduce quantum logics [1–3], suggesting that quantum mechanics requires a propositional calculus substantially different from all classical logics. He rigorously isolated a new algebraic structure for quantum logics, and studied its connections with quantum probabilities. Quantum and classical probabilities have points in common as well as differences. These differences and the properties of quantum probabilities have been intensively studied in the literature [4–14]. It is important to remark that not all authors believe that quantum probabilities are essentially of a different nature than those which arise in probability theory (see for example [15] for a recent account). Though this is a major question for probability theory and physics, it is not our aim in this work to settle this discussion.

There exist two important axiomatizations of classical probabilities. One of them was provided by Kolmogorov [16], a set theoretical approach based on Boolean sigma algebras of a sample space. Probabilities are defined as measures over subsets of a given set. Thus, the Kolmogorovian approach is set theoretical and usually identified (but not necessarily) with a frequentistic interpretation of probabilities. Some time later it was realized that quantum probabilities can be formulated as measures over non Boolean structures (instead of Boolean sigma algebras). This is the origin of the name "non-Boolean or non-Kolmogorovian" probabilities [8]. It is remarkable that the creation of quantum theory and the works on the foundations of probability by Kolmogorov were both developed at the same time, in the twenties.

An alternative approach to the Kolmogorovian construction of probabilities was developed by R.T. Cox [17,18]. Cox starts with a propositional calculus, intended to represent assertions which portray our knowledge about the world or system under investigation. As it is well known since the work of Boole [19], propositions of classical logic (CL) can be represented as a Boolean lattice, i.e., an algebraic structure endowed with lattice operations "$\wedge$", "$\vee$", and "$\neg$", which are intended to represent conjunction, disjunction, and negation, respectively, together with a partial order relation "$\leq$" which is intended to represent logical implication. Boolean lattices (as seen from an algebraic point of view) can be characterized by axioms [20–22]. By considering probabilities as an inferential calculus on a Boolean lattice, Cox showed that the axioms of classical probability can be deduced as a consequence of lattice symmetries, using entropy as a measure of information. Thus, differently form the set theoretical approach of Kolmogorov, the approach by Cox considers probabilities as an inferential calculus.

It was recently shown that Feynman's rules of quantum mechanics can be deduced from operational lattice structures using a variant of Cox's method [23,24,15,25,26] (see also [20,21]). For example, in [15,25] this is done by:

- first defining an operational propositional calculus on a quantum system under study, and after that,
- postulating that any quantum process (interpreted as a proposition in the operational propositional calculus) can be represented by a pair of real numbers and,
- using a variant of the method developed by Cox, showing that these pairs of real numbers obey the sum and product rules of complex numbers, and can then be interpreted as the quantum probability amplitudes which appear in Feynman's rules.

There is a long tradition with regards to the application of lattice theory to physics and many other disciplines. The quantum logical (QL) approach to quantum theory (and physics in general),

[1] By the term "quantum probabilities", we mean the probabilities that appear in quantum theory. As is well known, they are ruled by the well known formula $\mathrm{tr}(\rho P)$, where ρ is a density matrix representing a general quantum state and P is a projection operator representing an event (see Section 4 of this work for details).

initiated by von Neumann in [1], has been a traditional tool for studies on the foundations of quantum mechanics (see for example [27–37,5], and for a complete bibliography [6,38,39]).

The (QL) approach to physics bases itself on defining elementary tests and propositions for physical systems and then, studying the nature of these propositional structures. In some approaches, this is done in an operational way [29,40–44], and is susceptible to considerable generalization with arbitrary physical systems (not necessarily quantum ones). That is why the approach is also called *operational quantum logic* (OQL).[2] One of the most important goals of OQL is to impose operationally motivated axioms on a lattice structure in order that it can be made isomorphic to a projection lattice on a Hilbert space. There are different positions in the literature about the question of whether this goal has been achieved or not [14], and also, of course, alternative operational approaches to physics, such as the *convex operational* one [45–49,4]. In this work, we are interested in the great generality of the OQL approach. The operational approach presented in [4] bases itself only in the convex formulation of any statistical theory, and it can be shown that the more general structure which appears under reasonable operational considerations is a *σ-orthocomplemented orthomodular poset*, a more general class than *orthomodular lattices* (the ones which appear in quantum theory). We will come back to these issues and review the definitions for these structures below.

In this work we complement the work presented in [21,22] asking the following questions:

- Is it possible to generalize Cox's method to arbitrary lattices or more general algebraic structures?
- What happens if the Cox's method mentioned above is applied to general lattices (not necessarily distributive), representing general physical systems? And in particular, what happens if it is applied to the von Neumann's lattice of projection operators?
- Does the logical underlying structure of the theory determine the form and properties of the probabilities?

As we shall see below, it is possible to use these questions to give an alternative formulation of quantum probabilities. We will show that once the operational structure of the theory is fixed, the general properties of probability theory are – in a certain sense to be clarified below – determined. We also discuss the implications of our derivation for the foundations of quantum physics and probability theory, and compare with ours different approaches: the one presented in [15,25], the OQL approach, the operational approach of [4], and the traditional one (represented by the von Neumann formalism of Hilbertian quantum mechanics [50]).

The approach presented here shows itself to be susceptible to great generalization: we provide an algorithm for developing generalized probabilities using a combination of Cox's method with the OQL approach. This opens the door to the development of more general probability and information measures. This methodology is advantageous because, in the particular case of quantum mechanics, it includes mixed states in a natural way, unlike other approaches based only on pure states (like the ones presented in [26,23]).

The paper is organized as follows. In Section 2 we review the *QL* approach to physics as well as lattice theory. In Section 3 we revisit Kolmogorov's and Cox' approaches to probability. After that, in Section 4, we give a sketch concerning quantum probabilities and their differences with classical ones. In Section 5 we discuss the approach developed in [23,24,15,25,26]. In Section 6 we apply Cox's method for the formulation of non-Kolmogorovian probabilities using the algebraic properties of non-Boolean lattices and study several examples. Finally, in Section 7 some conclusions are drawn.

2. The lattice/operational approach to physics

The quantum logical approach to physics is vast and includes different programs. We will concentrate on the path followed by von Neumann and the operational approach developed by Jauch, Piron, and others. First, we recall the relationship between projection operators and elementary tests

[2] In this paper we will use the terms QL and OQL interchangeably, but it is important to remark that – though similar – they are different approaches.

in *QM*. After studying the examples of lattices applied to QM and CM, we review the main features of the *QL* approach. The reader familiar with these topics can skip this section.

2.1. Elementary notions of lattice theory

A partially ordered set (also called a poset) is a set X endowed with a partial ordering relation "$<$" satisfying

- 1—For all $x, y \in X$, $x < y$ and $y < x$ entail $x = y$.
- 2—For all $x, y, z \in X$, if $x < y$ and $y < z$, then $x < z$.

The notation "$x \leq y$" is used to denote "$x < y$" or "$x = y$". A lattice $\mathcal{L}$ will be a poset in which any two elements a and b have a unique supremum (the elements' least upper bound "$a \vee b$"; called their join) and an infimum (greatest lower bound "$a \wedge b$"; called their meet). Lattices can also be characterized as algebraic structures satisfying certain axiomatic identities imposed on operations "$\vee$" and "$\wedge$". For a *complete* lattice all its subsets have both a supremum (join) and an infimum (meet).

A *bounded* lattice has a greatest (or maximum) and least (or minimum) element, denoted 1 and 0 by convention (also called top and bottom, respectively). Any lattice can be converted into a bounded lattice by adding a greatest and least element, and every non-empty finite lattice is bounded. For any set A, the collection of all subsets of A (called the power set of A) can be ordered via subset inclusion to obtain a lattice bounded by A itself and the null set. Set intersection and union represent the operations meet and join, respectively.

Every complete lattice is a bounded lattice. While bounded lattice homomorphisms in general preserve only finite joins and meets, complete lattice homomorphisms are required to preserve arbitrary joins and meets. If P is a bounded poset, an orthocomplementation in P a unary operation "$\neg(\ldots)$" such that:

$$\neg(\neg(a)) = a \tag{1a}$$

$$a \leq b \longrightarrow \neg b \leq \neg a \tag{1b}$$

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

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

$$a \wedge \neg a = \mathbf{0} \tag{1d}$$

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 $a \leq \neg b$.

Distributive lattices are lattices for which the operations of join and meet are distributed over each other. A complete complemented lattice that is also distributive is a Boolean algebra. For a distributive lattice, the complement of x, when it exists, is unique. The prototypical examples of Boolean algebras are collections of sets for which the lattice operations can be given by set union and intersection, and lattice complementation by set theoretical complementation.

A *modular* lattice is one that satisfies the following self-dual condition (*modular law* or *modular identity*)

$$x \leq b \longrightarrow x \vee (a \wedge b) = (x \vee a) \wedge b. \tag{2}$$

Modular lattices arise naturally in algebra and in many other areas of mathematics. For example, the subspaces of a finite dimensional vector space form a modular lattice. Every distributive lattice is modular. In a not necessarily modular lattice, there may still be elements b for which the modular law holds in connection with arbitrary elements a and x ($\leq b$). Such an element is called a modular element. Even more generally, the modular law may hold for a fixed pair (a, b). Such a pair is called a modular pair, and there are various generalizations of modularity related to this notion and to semi-modularity.

An orthomodular lattice will be an ortholattice satisfying the orthomodular law:

$$x \leq b \longrightarrow x \vee (\neg x \wedge b) = b. \tag{3}$$

Orthomodularity is a weakening of modularity. As an example, the lattice $\mathcal{L}_{v.N}(\mathcal{H})$ of closed subspaces of a Hilbert space H (see Section 2.2) is orthomodular. $\mathcal{L}_{v.N}(\mathcal{H})$ is modular only if H is finite dimensional and strictly orthomodular for the infinite dimensional case.

The concept of a lattice's atom is of great physical importance. If $\mathcal{L}$ has a least 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 $0 < y < x$. One says that $\mathcal{L}$ is:

(i) *Atomic*, if for every nonzero element x of $\mathcal{L}$, there exists an atom a of $\mathcal{L}$ such that $a \leq x$.

(ii) Atomistic, if every element of $\mathcal{L}$ is a supremum of atoms.

2.2. Elementary measurements and projection operators

In QM, an elementary measurement given by a yes–no experiment (i.e., a test in which we get the answer "yes" or the answer "no"), is represented by a projection operator. If $\mathbb{R}$ is the real line, let $B(\mathbb{R})$ be the family of subsets of $\mathbb{R}$ such that

- 1—The family is closed under set theoretical complements.
- 2—The family is closed under denumerable unions.
- 3—The family includes all open intervals.

The elements of $B(\mathbb{R})$ are the *Borel subsets* of $\mathbb{R}$ [51]. Let $\mathcal{P}(\mathcal{H})$ be the set of all projection operators (or equivalently, the set of closed subspaces of $\mathcal{H}$). In QM, a projection valued measure (PVM) M, is a mapping

$$M : B(\mathbb{R}) \to \mathcal{P}(\mathcal{H}) \tag{4a}$$

such that

$$M(\emptyset) = 0 \tag{4b}$$

$$M(\mathbb{R}) = \mathbf{1} \tag{4c}$$

$$M(\cup_j(B_j)) = \sum_j M(B_j), \tag{4d}$$

for any disjoint denumerable family B_j. Also,

$$M(B^c) = \mathbf{1} - M(B) = (M(B))^{\perp}. \tag{4e}$$

Any elementary measurement is represented by a projection operator [50]. All operators representing observables can be expressed in terms of PVM's (and so, reduced to sets of elementary measurements), via the spectral decomposition theorem, which asserts that the set of spectral measurements may be put in a bijective correspondence with the set $\mathcal{A}$ of self adjoint operators of $\mathcal{H}$ [51].

The set of closed subspaces $\mathcal{P}(\mathcal{H})$ of any quantum system can be endowed with a lattice structure: $\mathcal{L}_{v.N}(\mathcal{H}) = \langle \mathcal{P}(\mathcal{H}), \leq, \wedge, \vee, \neg, 0, 1\rangle$, where "$\leq$" is the set theoretical inclusion "$\subseteq$", "$\wedge$" is set theoretical intersection "$\cap$", "$\vee$" is the closure of the sum "$\oplus$", 0 is the empty set $\emptyset$, 1 is the total space $\mathcal{H}$ and $\neg(S)$ is the orthogonal complement of a subspace S [7]. Closed subspaces can be put in one to one correspondence with projection operators. *Thus, elementary tests in QM, which are represented by projection operators, can be endowed with a lattice structure*. This lattice was called "Quantum Logic" by Birkhoff and von Neumann [1]. We will refer to this lattice as the *von Neumann-lattice* ($\mathcal{L}_{v.N}(\mathcal{H})$) [7].

The analogous of this structure in Classical Mechanics (*CM*) was provided by Birkhoff and von Neumann [1]. Take for example the following operational propositions on a classical harmonic oscillator: "the energy is equal to E_0" and "the energy is lesser or equal than E_0". The first one corresponds to an ellipse in phase space, and the second to the ellipse and its interior. This simple example shows that operational propositions in *CM* can be represented by subsets of the phase space. Thus, given a classical system S with phase space Γ, let $\mathcal{P}(\Gamma)$ represent the set formed by all the subsets of Γ. This set can be endowed with a lattice structure as follows. If "$\vee$" is represented by set union, "$\wedge$" by set intersection, "$\neg$" by set complement (with respect to Γ), $\leq$ is represented by set inclusion, and $\mathbf{0}$ and $\mathbf{1}$ are represented by $\emptyset$ and Γ respectively, then $\langle \mathcal{P}(\Gamma), \leq, \wedge, \vee, \neg, 0, 1\rangle$ conform a complete bounded lattice. This is the lattice of propositions of a classical system, which as

it is well known, *is a Boolean one*. Thus, $\mathcal{P}(\Gamma)$, as well as $\mathcal{P}(\mathcal{H})$, can be endowed with a propositional lattice structure.

2.3. The quantum logical approach to physics

We have seen that operational propositions of quantum and classical systems can be endowed with lattice structures. These lattices were Boolean for classical systems, and non distributive for quantum ones. This fact, discovered by von Neumann [1], raised a lot of interesting questions. The first one is: is it possible to obtain the formalism of *QM* (as well as *CM*) by imposing suitable axioms on a lattice structure? The surprising answer is *yes, it is possible*. But the road which led to this result was fairly difficult and full of obstacles. In the first place, it was a very difficult mathematical task to demonstrate that a suitably chosen set of axioms on a lattice would yield a representation theorem which would allow one to recover Hilbertian *QM*. The first result was obtained by Piron, and the final demonstration was given by Solèr in 1995 [52] (see also [6, page 72]). One of the advantages ascribed to this approach was that the axioms imposed on a lattice structure could be given a clear operational interpretation: unlike the Hilbert space formulation, whose axioms have the disadvantage of being *ad hoc* and physically unmotivated, the quantum logical approach would be clearer and more intuitive from a physical point of view. But of course, the operational validity of the axioms imposed on the lattice structure was criticized by many authors (as an example, see [14]).

The second important question raised by the von Neumann discovery was: given that *QM* and *CM* can be described by operational lattices, is it possible to formulate the entire apparatus of physics in lattice theoretical terms? Given *any* physical system, quantum, classical, or obeying more general tenets, it is always possible to define an operational propositional structure on it using the notion of elementary tests. A very general approach to physics can be given using *event structures*, which are sets of events endowed with probability measures satisfying certain axioms [4]. It can be shown (see [4, Chapter 3]) that any event structure is isomorphic to a *σ-orthocomplete orthomodular poset*, which is an orthocomplemented poset $\mathcal{P}$, satisfying the orthomodular identity (3), and for which if $a_i \in \mathcal{P}$ and $a_i \perp a_j$ $(i \neq j)$, this implies that $\bigvee a_i$ exists for $i = 1, 2, \ldots$. Remark that event structures (or σ-orthocomplete orthomodular posets) need not be lattices. However, lattices are very general structures and encompass most important examples. Consequently, we will work with orthomodular lattices in this paper (and indicate which results can be easily extended to σ-orthocomplete orthomodular posets).

There are other general approaches to statistical theories. One of them is the convex operational one [53–56], which consists of imposing axioms on a convex structure (formed by physical states). Indeed, the convex operational approach is even more general than the quantum logical, but we will not discuss this issue in detail here (although a link with it will be discussed in Section 6.5).

3. Cox vs. Kolmogorov

In this section we will review two different approaches to probability theory. On one hand, the Cox's approach, in which probabilities are considered as measures of the plausibility of a given event or happening. On the other hand, the traditional Kolmogorovian one, a set theoretical approach which is compatible with the interpretation of probabilities as frequencies.

3.1. Kolmogorov

Given a set Ω, let us consider a σ-algebra Σ of Ω. Then, a probability measure will be given by a function μ such that

$$\mu : \Sigma \to [0, 1] \tag{5a}$$

which satisfies

$$\mu(\emptyset) = 0 \tag{5b}$$

$$\mu(A^c) = 1 - \mu(A), \tag{5c}$$

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

$$\mu\left(\bigcup_{i\in I} A_i\right) = \sum_i \mu(A_i) \tag{5d}$$

where conditions (5) are the well known axioms of Kolmogorov. The triad (Ω, Σ, μ) is called a *probability space*. Depending on the context, probability spaces obeying Eqs. (5) are usually referred as Kolmogorovian, classical, commutative or Boolean probabilities [4].

It is possible to show that if (Ω, Σ, μ) is a Kolmogorovian probability space, the *inclusion–exclusion principle* holds

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \tag{6}$$

or (as expressed in logical terms)

$$\mu(A \vee B) = \mu(A) + \mu(B) - \mu(A \wedge B). \tag{7}$$

As remarked in [57], Eq. (6) 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 each event occurs in a series of N repetitions, then (6) trivially holds.

As we shall discuss below, this principle no longer holds in QM, a fact linked to the non-Boolean QM-character. Thus, the relative-frequencies' interpretation of quantum probabilities becomes problematic [57]. The QM example shows that non-distributive propositional structures play an important role in probability theories *different from* that of Kolmogorov.

3.2. Cox's approach

Propositions of classical logic can be endowed with a Boolean lattice structure [19]. The logical implication "$\longrightarrow$" is associated with a partial order relation "$\leq$", the conjunction "and" with the greatest lower bound "$\wedge$", disjunction "or" with the lowest upper bound "$\vee$", and negation "not" is associated with complement "$\neg$". Boolean lattices can be characterized as ortholattices satisfying:

- L1. $x \vee x = x, x \wedge x = x$ (idempotence).
- L2. $x \vee y = y \vee x, x \wedge y = y \wedge x$ (commutativity).
- L3. $x \vee (y \vee z) = (x \vee y) \vee z, x \wedge (y \wedge z) = (x \wedge y) \wedge z$ (associativity).
- L4. $x \vee (x \wedge y) = x \wedge (x \vee y) = x$ (absorption).
- D1. $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ (distributivity 1).
- D2. $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ (distributivity 2).

It is well known that Boolean lattices can be represented as subsets of a given set, with "$\leq$" represented as set theoretical inclusion $\subseteq$, "$\vee$" represented as set theoretical union "$\cup$", "$\wedge$" represented as set intersection "$\cap$", and $\neg$ represented as the set theoretical complement "$(\ldots)^c$".

As a typical feature, Cox develops classical probability theory as an inferential calculus on Boolean lattices. A real valued function φ representing the degree to which a proposition y implies another proposition x is postulated, and its properties deduced from the algebraic properties of the Boolean lattice (Eqs. (1) and (3.2)). These algebraic properties define functional equations [58] which determine the possible elections of φ up to rescaling. It turns out that $\varphi(x|y)$ – if suitably normalized – satisfies all the properties of a Kolmogorovian probability (Eqs. (5)). The deduction will be omitted here, and the reader is referred to [17,18,20,21,24] for detailed expositions.

Despite their formal equivalence, there is a great conceptual difference between the approaches of Kolmogorov and Cox. In the Kolmogorovian approach probabilities are naturally interpreted (but not necessarily) as relative frequencies in a sample space. On the other hand, the approach developed by Cox, considers probabilities as a measure of the degree of belief of an intelligent agent, on the truth of proposition x if it is known that y is true. This measure is given by the real number $\varphi(x|y)$, and in this way the Cox's approach is more compatible with a *Bayesian* interpretation of probability theory.

4. Quantum vs. classical probabilities

In this section we will introduce quantum probabilities and look at their differences with classical ones. A great part of the hardship faced by Birkhoff and von Neumann in developing the logic of quantum mechanics was due to the inadequacies of classical probability theory. Their point of view was that any statistical physical theory could be regarded as a probability theory, founded on a calculus of events. These events should be the experimentally verifiable propositions of the theory, and the structure of this calculus was to be deduced from empirical considerations, which, for the quantum case, resulted in an orthomodular lattice [1,11]. We remark on the great generality of this conception: there is no need to restrict it to physics. Any statistical theory formulated as an event structure fits into this scheme.

In the formulation of both classical and quantum statistical theories, states can be regarded as representing consistent probability assignments [9,53]. In the quantum mechanics instance *this "states as mappings" visualization* is achieved via *postulating* a function [7]

$$s : \mathcal{P}(\mathcal{H}) \to [0; 1] \tag{8a}$$

such that:

$$s(\mathbf{0}) = 0 \quad (\mathbf{0} \text{ is the null subspace}). \tag{8b}$$

$$s(P^{\perp}) = 1 - s(P), \tag{8c}$$

and, for a denumerable and pairwise orthogonal family of projections P_j

$$s\left(\sum_j P_j\right) = \sum_j s(P_j). \tag{8d}$$

Gleason's theorem [59,60], tells us that if the dimension of $\mathcal{H} \geq 3$, any measure s satisfying (8) can be put in correspondence with a trace class operator (of trace one) ρ_s via the correspondence:

$$s(P) := \mathrm{tr}(\rho_s P). \tag{9}$$

And vice versa: using Eq. (9) any trace class operator of trace one defines a measure as in (8). Thus, Eqs. (8) define a probability: to any elementary test (or event), represented by a projection operator P, $s(P)$ gives us the probability that the event P occurs, and this is experimentally granted by the validity of Born's rule. But in fact, (8) is not a classical probability, because it does not obey Kolmogorov's axioms (5). The main difference comes from the fact that the σ-algebra in (5) is Boolean, while $\mathcal{P}(\mathcal{H})$ is not. Thus, quantum probabilities are also called non-Kolmogorovian (or non-Boolean) probability measures. The crucial fact is that, in the quantum case, *we do not have a σ-algebra, but an orthomodular lattice of projections.*

One of the most important ensuing differences expresses itself in the fact that Eq. (6) is no longer valid in QM. Indeed, it may happen that

$$s(A) + s(B) \leq s(A \vee B) \tag{10}$$

for A and B suitably chosen elementary sharp tests (see [4, Chapter 2]). Another important difference comes from the difficulties which appear when one tries to define a quantum conditional probability (see for example [4,8] for a comparison between classical and quantum probabilities). Quantum probabilities may also be considered as a generalization of classical probability theory: 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* [4,8].

We are thus faced with the following fact: on the one hand, there exists a generalization of classical probability theory to non-Boolean operational structures. On the other hand, Cox derives classical probabilities from the algebraic properties of classical logic. As we shall see in detail below, this readily implies that probabilities in CM are determined by the operational structure of classical propositions (given by subsets of phase space). The question is: is it possible to generalize Cox's method to arbitrary propositional structures (representing the operational propositions of an arbitrary theory) even when

they are not Boolean? What would we expect to find? We will see that the answer to the first question is *yes*, and for the second, it is reasonable to recover quantum probabilities (Eq. (8)). This approach may serve as a solution for a problem posed by von Neumann. In his words:

> "In order to have probability all you need is a concept of all angles, I mean, other than 90. Now it is perfectly true that in geometry, as soon as you can define the right angle, you can define all angles. Another way to put it is that if you take the case of an orthogonal space, those mappings of this space on itself, which leave orthogonality intact, leave all angles intact, in other words, in those systems which can be used as models of the logical background for quantum theory, it is true that as soon as all the ordinary concepts of logic are fixed under some isomorphic transformation, all of probability theory is already fixed.... This means however, that one has a formal mechanism in which, logics and probability theory arise simultaneously and are derived simultaneously [57]".

and, as remarked by M. Redei [57]:

> "It was simultaneous emergence and mutual determination of probability and logic what von Neumann found intriguing and not at all well understood. He very much wanted to have a detailed axiomatic study of this phenomenon because he hoped that it would shed "...a great deal of new light on logics and probability alter the whole formal structure of logics considerably, if one succeeds in deriving this system from first principles, in other words from a suitable set of axioms". (quote) He emphasized – and this was his last thought in his address – that it was an entirely open problem whether/how such an axiomatic derivation can be carried out".

The problem posed above has remained thus far unanswered, and this work may be considered as concrete step towards its solution. Before entering the subject, let us first review an alternative approach.

5. Alternative derivation of Feynman's rules

Refs. [15,25,26] present a novel derivation of Feynman's rules for quantum mechanics, based on a modern reformulation [24] of Cox's ideas on the foundations of probability [17,18]. To start with, an experimental logic of processes is defined for quantum systems. This is done in such a way that the resulting algebra is a distributive one. Given n measurements $M_1,\ldots, M_n$ on a given system, with results $m_1, m_2, \ldots, m_n$, the later are organized in a *measuring sequence* $A = [m_1, m_2, \ldots, m_n]$ as a particular process. The measuring sequence $A = [m_1, m_2, \ldots, m_n]$ must not be confused with the conditional (logical) proposition of the form $(m_2, \ldots, m_n|m_1)$. Sequence A has associated a probability $P(A) = Pr(m_n, \ldots, m_2|m_1)$ of obtaining outcomes $m_2, \ldots, m_n$ conditional upon obtaining m_1 [15].

If each of the m_i's has two possible values, 1 and 2, a measuring sequence of three measurements is for example $A_1 = [1, 2, 1]$. Another one could be $A_2 = [1, 1, 2]$, and so on.

As is explained in Ref. [15]:

> "A particular outcome of a measurement is either atomic or coarse-grained. An atomic outcome cannot be more finely divided in the sense that the detector whose output corresponds to the outcome cannot be sub-divided into smaller detectors whose outputs correspond to two or more outcomes. A coarse-grained outcome is one that does not differentiate between two or more outcomes".

Thus, if we want to "coarse grain" a certain measurement, say M_2, we can unite the two outcomes in a joint outcome $(1, 2)$, yielding the experiment (measurement) $\tilde{M}_2$. Thus, a possible sequence obtained by the replacement of M_2 by $\tilde{M}_2$ could be $[1, (1, 2), 1]$. This is used to define a logical operation

$$[m_1, \ldots, (m_i, m_i'), \ldots, m_n] = [m_1, \ldots, m_i, \ldots, m_n] \vee [m_1, \ldots, m_i', \ldots, m_n]. \tag{11}$$

It is intended that sequences of measurements can be compounded. For example, if we have $[m_1, m_2]$ and $[m_2, m_3]$, we have also the sequence $[m_1, m_2, m_3]$, paving the way for the general definition

$$[m_1, \ldots, m_j, \ldots, m_n] = [m_1, \ldots, m_j] \cdot [m_j, \ldots, m_n]. \tag{12}$$

Given measuring sequences A, B and C, these operations satisfy

$$A \vee B = B \vee A \tag{13a}$$
$$(A \vee B) \vee C = A \vee (B \vee C) \tag{13b}$$
$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \tag{13c}$$
$$(A \vee B) \cdot C = (A \cdot C) \vee (B \cdot C) \tag{13d}$$
$$C \cdot (A \vee B) = (C \cdot A) \vee (C \cdot B), \tag{13e}$$

and thus, we have commutativity and associativity of the operation "$\vee$", associativity of the operation "$\cdot$", and right- and left-distributivity of "$\cdot$" over "$\vee$".

We had already seen in Section 3 that the method of Cox consists of deriving probability and entropy from the symmetries of a Boolean lattice, intended to represent our propositions about the world, while probability is interpreted as a measure of knowledge about an inference calculus. Once Eqs. (13) are cast, the setup for the derivation of Feynman's rules is ready. The path to follow now is to apply Cox's method to the symmetries defined by Eqs. (13), but this cannot be done straightforwardly. In order to proceed, an important assumption has to be made: each measuring sequence will be represented by a pair of real numbers. This – non operational – assumption is justified in [15] using Bohr's complementarity principle. As we shall see below, the method proposed in this article is an alternative one, which is more direct and systematic, and makes the introduction of these assumptions somewhat clearer.

Once a pair of real numbers is assigned to any measuring sequence, the authors of [15] reasonably assume that Eqs. (13) induce operations onto pairs of reals numbers. If measuring sequences A, B, etc. induce pairs of real numbers $\mathbf{a}$, $\mathbf{b}$, etc., then, we should have

$$\mathbf{a} \vee \mathbf{b} = \mathbf{b} \vee \mathbf{a} \tag{14a}$$
$$(\mathbf{a} \vee \mathbf{b}) \vee \mathbf{c} = \mathbf{a} \vee (\mathbf{b} \vee \mathbf{c}) \tag{14b}$$
$$(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c}) \tag{14c}$$
$$(\mathbf{a} \vee \mathbf{b}) \cdot \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) \vee (\mathbf{b} \cdot \mathbf{c}) \tag{14d}$$
$$\mathbf{c} \cdot (\mathbf{a} \vee \mathbf{b}) = (\mathbf{c} \cdot \mathbf{a}) \vee (\mathbf{c} \cdot \mathbf{b}). \tag{14e}$$

We easily recognize in (14) operations satisfied by the complex numbers' field (provided that the operations are interpreted as sum and product of complex numbers). If they constituted the only possible instance, sequences represented by pairs of real numbers would be complex numbers, and thus, we could easily have Feynman's rules. However, complex numbers are not the only entities that satisfy (14). There are other such entities, and thus, extra assumptions have to be made in order to restrict possibilities. These additional assumptions are presented in [15,25], and improved upon in [26]. We list them below (and refer the reader to Refs. [15,25,26] for details).

- Pair symmetry.
- Additivity condition.
- Symmetric bias condition.

Leaving aside the fact that these extra assumptions are more or less reasonable (justifications for their use are given in [26]), it is clear that the derivation is quite indirect: the experimental logic is thus defined in order to yield algebraic rules compatible with complex multiplication (and the rest of the strategy is to make further assumptions in order to discard other fields different from that of complex numbers). Further, the experimental logic characterized by Eqs. (13) is not the only possibility, as we have seen in Section 2.

In the rest of this work, we will apply Cox's method to general propositional structures according to the quantum logical approach. We will see that this allows for a new perspective which sheds light onto the structure of non-Boolean probabilities, and is at the same time susceptible to great generalization. It opens the door to a general derivation of alternative kinds of probabilities, including quantum and classical theories as particular cases.

Yet another important remark is in order. As noted in the Introduction, the work presented in [15,24,25] – as well as ours – is a combination of two approaches: (1) the one which defines

propositions in an empirical way (something which it shares with the OQL approach) and (2) that of Cox. Cox's spirit was to derive probabilities out of Chomsky's generative propositional structures that are ingrained in our brain [61], and this Boolean structure is independent of any experimental information. This does not imply, though, that the empirical logic needs to satisfy the same algebra than pervades our thinking, and that is indeed what happens. In this sense, any derivation involving empirical or operational logics deviates from the original intent of Cox. As we shall see, this is not a problem, but rather an important advantage in practice.

6. Cox's method applied to non-Boolean algebras

As seen in Section 2, *operationally motivated axioms imposed on a lattice's propositional structure can be used to describe quantum mechanics and other theories as well.* Disregarding the discussion about the operational validity of this construction, we are only interested in the fact that the embodiment is feasible. Similar constructions can be made for many physical systems, beyond quantum mechanics: the connection between any theory and experience is given by an event structure (elementary tests), and these events can be organized in a lattice structure in most examples of interest.

Thus, our point of departure will be the fact that physical systems can be represented by propositional lattices, and that these lattices need not be necessarily distributive. We will consider atomic orthomodular lattices. Given a system S, and its propositional lattice $\mathcal{L}$, we proceed to apply Cox's method in order to develop an inferential calculus on $\mathcal{L}$.

6.1. Classical mechanics

We start with classical mechanics (that theory satisfying Hamilton's equations). Given a classical system S_C, the propositional structure is a Boolean one, isomorphic to a perfectly Boolean lattice used in our logical language (i.e., regarding its algebraic structure, it is the same as the one used by Cox). Accordingly, as shown in Section 3 (proceeding in the same way as Cox [18,17]), the corresponding probability calculus has to be the one which obeys the laws of Kolmogorov (it satisfies – in particular – Eqs. (5)), and the corresponding information measure is Shannon's, as expected.

6.2. Quantum case

As shown by Birkhoff and von Neumann in [1], if we follow the above path and try to define the propositional structure for a quantum system S_Q we find an orthomodular lattice $\mathcal{L}_{v.N}(\mathcal{H})$ isomorphic to the lattice of projections $\mathcal{P}(\mathcal{H})$. What are we going to find if we apply instead Cox's method? It stands to reason that we would encounter a non-Boolean probability measure with the properties *postulated* in Section 4 (Eqs. (8)). Let us see that this is indeed the case.

The first thing to remark is that in this derivation we assume to have a non-Boolean lattice $\mathcal{L}_{v.N}(\mathcal{H})$, isomorphic to the lattice of projections $\mathcal{P}(\mathcal{H})$. We must show that the "degree of implication" measure $s(\cdots)$ demanded by Cox's method satisfies Eqs. (8). We will only consider the case of prior probabilities. This means that we ask for the probability that a certain event happens for a given state of affairs, i.e., a concrete preparation of the system under certain circumstances (which could be natural or artificial). Thus, we are looking for a function to the real numbers s such that it is non-negative and $s(P) \leq s(Q)$ whenever $P \leq Q$.

Under these assumptions, let us consider the operation "$\vee$". As the direct sum of subspaces is associative, "$\vee$" will be associative too. If P and Q are orthogonal projections ($P \perp Q$), then $P \wedge Q = \mathbf{0}$ (otherwise, there would be a vector in P which is not orthogonal to every vector of Q). Next, we consider the relationship between $s(P)$, $s(Q)$, and $s(P \vee Q)$. As $P \wedge Q = \mathbf{0}$, it should happen that

$$s(P \vee Q) = F(s(P), s(Q)), \tag{15}$$

with F a function to be determined. Add now a third proposition R (notice that, to do this, we need a space of dimension $d \geq 3$, an interesting analogy with Gleason's theorem), such that $P \perp R$, $Q \perp R$, and $Q \perp P$ (and thus $P \wedge R = \mathbf{0}$, $Q \wedge R = \mathbf{0}$, and $Q \wedge P = \mathbf{0}$). Build now the element $(P \vee Q) \vee R$.

Then, because of the associativity of "$\vee$", we arrive at the following result

$$s((P \vee Q) \vee R) = s(P \vee (Q \vee R)), \tag{16}$$

and thus (using (15)),

$$F(F(s(P), s(Q)), s(R)) = F(s(P), F(s(Q), s(R))). \tag{17}$$

The algebraic properties of associativity for $\vee$ and $\perp$ are the only prerequisite for this result. Thus, proceeding as in [20,21,24] (and using the solutions to functional equations of the form (17) studied in [58]), we have that — up to a re-scaling:

$$s(P \vee Q) = s(P) + s(Q) \tag{18}$$

whenever $P \perp Q$. It thus follows that for any finite family of orthogonal projections P_j, $1 \leq j \leq n$, we have $s(P_1 \vee P_2 \vee \cdots \vee P_n) = s(P_1) + s(P_2) + \cdots + s(P_n)$. Now, as any projection P satisfies $P \leq \mathbf{1}$, then $s(P) \leq s(\mathbf{1})$, and we can assume without loss of generality the normalization condition $s(\mathbf{1}) = 1$. Thus, for any denumerable pairwise orthogonal infinite family of projections P_j, we have for each n

$$\sum_{j=1}^{n} s(P_j) = s\left(\bigvee_{j=1}^{n} P_j\right) \leq 1. \tag{19}$$

As $s(P_j) \geq 0$ for each j, the sequence $s_n = s(\bigvee_{j=1}^{n} P_j)$ is monotone, bounded from above, and thus converges. We write then

$$s\left(\bigvee_{j=1}^{\infty} P_j\right) = \sum_{j=1}^{\infty} s(P_j), \tag{20}$$

and we recover condition (8d) of the axioms of quantum probability. Now, given any proposition $\mathcal{L}_{v.N}(\mathcal{H})$, consider $P^{\perp}$. As $P \vee P^{\perp} = \mathbf{1}$, and P is orthogonal to $P^{\perp}$, we have

$$s(P \vee P^{\perp}) = s(P) + s(P^{\perp}) = s(\mathbf{1}) = 1. \tag{21}$$

In other words,

$$s(P^{\perp}) = 1 - s(P), \tag{22}$$

which is nothing but condition (8c). 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 (8b).

This section shows that using the algebraic properties of $\mathcal{L}_{v.N}$, it is possible to derive the form of the quantum probabilities which, in the light of this discussion, do not need to be postulated.

Thus, we have proved that s is a probability measure on $\mathcal{L}_{v.N}$. Is there any possibility that s differs from the standard formulation of a quantum probability measure as a density matrix using Born's rule? The answer is no, and this is granted by Gleason's theorem, because we have proved that s satisfies Eqs. (8), and Gleason's theorem leaves no alternative (if the dimension of $\mathcal{H} \geq 3$).

An important question is the following: which will be the effect of non-distributivity? As we saw in Section 4, classical probabilities are sub-additive, i.e., they satisfy

$$\mu(A \vee B) \leq \mu(A) + \mu(B), \tag{23}$$

and this is linked to the stronger assertion of Eq. (6) (see also [7, page 104]). But as we have seen, it is indeed the case that the analogous of Eq. (23) does not hold for quantum probabilities.

We show below that this derivation is susceptible to generalization. Indeed, the derivation relies mainly on the algebraic properties of the lattice of projections, i.e., in its *non-distributive* lattice structure.

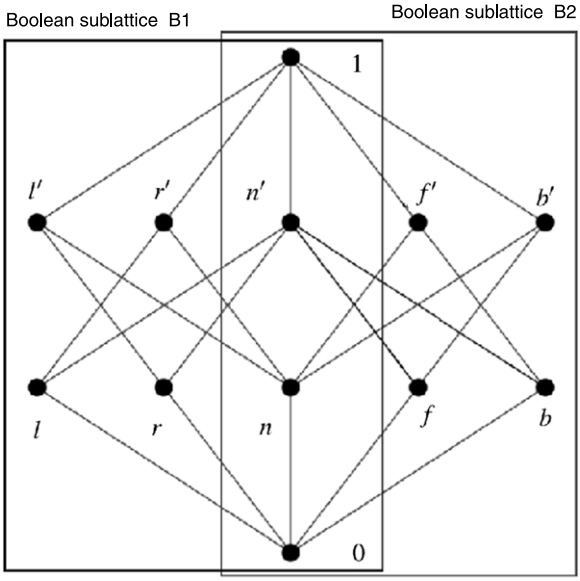

Fig. 1. Hasse diagram for L_{12}. The Boolean sublattices B1 and B2 corresponding to the two complementary contexts are indicated by squared lines.

6.3. A finite non-distributive example

There are many systems of interest which can be represented by finite lattices. Many toy models serve to illustrate special features of different theories. Let us start first by analyzing L_{12}, a non distributive lattice which may be considered as the union of two incompatible experiments [62]. The Hasse diagram of L_{12} is represented in Fig. 1. An example of L_{12} is provided by a firefly which flies inside a room. The first experiment is to test if the firefly shines on the right side (r) of the room, or on the left (l), or if it does not shine at all (n). Another experiment consists of testing if the firefly shines at the front of the room (f), or at the back (b) of it, or if it does not shine (n). It is forbidden to conduct both experiments at the same time (and thus, emulating contextuality).

Applying the Cox's method to the Boolean sublattices of L_{12} (see Fig. 1) and suitably normalizing, we obtain classical probabilities for each one of them. It is also easy to find that $P(l) + P(r) + P(f) + P(b) + P(n) = 2 - P(n)$, a quantity which may be greater than 1. The last equation implies explicitly that the exclusion–inclusion law (6) does not hold. This is due to the global non-distributivity of L_{12}. It is also easy to show that $P(l) + P(r) = P(f) + P(b)$, yielding a non trivial relationship between atoms (something which does not occur in a distributive lattice).

6.4. General derivation

Let $\mathcal{L}$ be an atomic orthomodular lattice. We will compute prior probabilities, i.e., we will assume that $\mathcal{L}$ represents the propositional structure of a given system – physical or not –, and that we want to ascertain how likely a given event is when represented by a lattice element $a \in \mathcal{L}$. One assumes that the system has undergone a preparation, i.e., we take it that there exists a *state of affairs*.

We must define a function $s : \mathcal{L} \longrightarrow \mathbb{R}$, such that it is always non-negative

$$s(a) \geq 0 \quad \forall a \in \mathcal{L} \tag{24a}$$

and is also order preserving

$$a \leq b \longrightarrow s(a) \leq s(b). \tag{24b}$$

We will show that under these rather general assumptions a probability theory can be developed. The order preserving assumption readily implies that $s(a) \leq s(\mathbf{1})$ for all $a \in \mathcal{L}$. We will also assume that $s(\mathbf{1}) = K$, a finite real number.

Now, as an ortholattice is complemented (Eqs. (1)), we will always have that $\neg\neg a = a$ for all $a \in \mathcal{L}$. Accordingly,

$$s(\neg\neg a) = s(a), \tag{25}$$

for all a. Next, it is also reasonable to assume that $s(\neg a)$ is a function of $s(\neg a)$, say $s(\neg a) = g(s(a))$. Thus, Eqs. (25) and (1a) imply

$$g(g(s(a))) = s(a), \tag{26}$$

or, in other words,

$$g(g(x)) = x, \tag{27}$$

for positive x. A family of functions which satisfy (26) are $g(x) = x$ and $g(x) = c - x$, where c is a real constant.[3] We discard the first possibility because if true, we would have $s(\mathbf{0}) = s(\neg\mathbf{1}) = g(s(\mathbf{1})) = s(\mathbf{1})$. But if $s(\mathbf{0}) = s(\mathbf{1})$, because of $\mathbf{0} \leq x \leq \mathbf{1}$ for all $x \in \mathcal{L}$, we have $s(\mathbf{0}) = s(x) = s(\mathbf{1})$, and our measure would be trivial. Thus, the only non-trivial option – up to rescaling – is $s(\neg a) = c - s(a)$.

Now, let us see what happens with the "$\vee$" operation. As $\mathcal{L}$ is orthocomplemented, the orthogonality notion for elements is available (see Section 2.1). If $a, b \in \mathcal{L}$ and $a \perp b$, because of (1b), we have that $a \wedge b = \mathbf{0}$. Thus, it is reasonable to assume that $s(a \vee b)$ is a function of $s(a)$ and $s(b)$ only, i.e., $s(a \vee b) = f(s(a), s(b))$. By associativity of the "$\vee$" operation, $(a \vee b) \vee c = a \vee (b \vee c)$ for any $a, b, c \in \mathcal{L}$, and this implies then that $s((a \vee b) \vee c) = s(a \vee (b \vee c))$. If a, b, and c are orthogonal, we will have for the left hand side $s((a \vee b) \vee c) = f(f(s(a), s(b)), s(c))$ and $s(a \vee (b \vee c)) = f(s(a), f(s(b), s(c)))$ for the right hand side. Thus,

$$f(f(s(a), s(b)), s(c)) = f(s(a), f(s(b), s(c))), \tag{28}$$

or, in a simpler form

$$f(f(x, y), z) = f(x, f(y, z)), \tag{29}$$

with x, y, and z positive real numbers. As shown in [58], the only solution (up to re-scaling) of (29) is $f(x, y) = x + y$. We have thus shown that if $a \perp b$

$$s(a \vee b) = s(a) + s(b), \tag{30}$$

and we will also have

$$s(a_1 \vee a_2 \cdots \vee a_n) = s(a_1) + s(b_2) + \cdots + s(a_n), \tag{31}$$

whenever $a_1, a_2, \ldots, a_n$ are pairwise orthogonal. Suppose now that $\{a_i\}_{i \in \mathbb{N}}$ is a family of pairwise orthogonal elements of $\mathcal{L}$. For any finite n, we have that $a_1 \vee a_2 \vee \cdots \vee a_n \leq \mathbf{1}$, and thus $s(a_1 \vee a_2 \vee \cdots \vee a_n) = s(a_1) + s(b_2) + \cdots + s(a_n) \leq s(\mathbf{1}) = K$. Then, $s_n = s(a_1 \vee a_2 \vee \cdots \vee a_n)$ is a monotone sequence bounded from above, and thus it converges to a real number. As $\bigvee\{a_i\}_{i \in \mathbb{N}} = \lim_{n \longrightarrow \infty} \bigvee_{i=1}^{n} a_i$, we can write

$$s\left(\bigvee\{a_i\}_{i \in \mathbb{N}}\right) = \sum_{i=1}^{\infty} s(a_i). \tag{32}$$

In any orthomodular lattice we have $\mathbf{1} \perp \mathbf{0}$ (because $\mathbf{0} \leq \neg\mathbf{1} = \mathbf{0}$), and $\mathbf{1} \vee \mathbf{0} = \mathbf{1}$. Thus, $s(\mathbf{1} \vee \mathbf{0}) = s(\mathbf{1}) = s(\mathbf{1}) + s(\mathbf{0})$. Accordingly, $s(\mathbf{0}) = 0$. As $\neg\mathbf{1} = \mathbf{0}$, $s(\mathbf{0}) = c - s(\mathbf{1})$. Thus, $s(\mathbf{1}) = c$ and then, $c = K$. We will not lose generality if we assume the normalization condition $K = 1$.

[3] There are additional solutions to this equation, but, if we suitably choose scales, we can disregard them for these two cases. For a discussion on re-scaling of measures, their meaning and validity, see [18].

The results of this section show that in *any* orthomodular lattice, a reasonable measure s of plausibility of a given event must satisfy that, for any orthogonal denumerable family $\{a_i\}_{i\in\mathbb{N}}$, one has (up to rescaling)

$$s\left(\bigvee\{a_i\}_{i\in\mathbb{N}}\right) = \sum_{i=1}^{\infty} s(a_i) \tag{33a}$$

$$s(\neg a) = 1 - s(a) \tag{33b}$$

$$s(\mathbf{0}) = 0. \tag{33c}$$

Why do Eqs. (33) define non-classical (non-Kolmogorovian) probability measures? In a non-distributive orthomodular lattice there always exist elements a and b such that

$$(a \wedge b) \vee (a \wedge \neg b) < a, \tag{34}$$

so that (using $(a \wedge \neg b) \perp (a \wedge b)$), $s((a \wedge \neg b) \vee (a \wedge b)) = s(a \wedge \neg b) + s(a \wedge b) \leq s(a)$. The inequality can be strict, as the quantum case shows. But in any classical probability theory, by virtue of the inclusion–exclusion principle (Eq. (7)), we always have $s(a \wedge \neg b) + s(a \wedge b) = s(a)$. This simple fact shows that our measures will be non-classical in the general case.

In this way, we provide an answer to the problem posed by von Neumann and discussed in Section 4: *the algebraic and logical properties of the operational event structure determine up to rescaling the general form of the probability measures which can be defined over the lattice.* Accordingly, we did present a generalization of the Cox method to non-Boolean structures, namely orthomodular lattices, and then *we have indeed deduced a generalized probability theory.*

Remark that Boolean lattices are also orthomodular. This means that our derivation is a generalization of that of Cox, and when we face a Boolean structure, classical probability theory will be deduced exactly as in [17,18].

Another important remark is that – as the reader may check – the derivation presented here *is also valid* for σ-orthocomplete orthomodular posets (see Section 2 of this work), and thus, of great generality. Indeed, in these structures the disjunction exists for denumerable collections of orthogonal elements and it is associative. This fact, together with orthocomplementation, make the above results remain valid in these structures, but as mentioned in Section 2, σ-orthocomplete orthomodular posets are not lattices in the general case. In this way, we remark that the generalization of the Cox's method goes well beyond lattice theory. Thus, yielding a generalized probability calculus which includes the general physical approach presented for example, in [4].

6.5. A general methodology

At this stage it is easy to envisage just how a general method could be developed. One first starts by identifying the algebraic structure of the elementary tests of a given theory $\mathcal{T}$. They determine all observable quantities and, in the most important examples, they are endowed with a lattice structure. Once the algebraic properties of the pertinent lattice are fixed, Cox's method can be used to *determine* (up to rescaling) the general properties of prior probabilities. Of course, it does not predetermine the particular values that these probabilities might take on the atoms of the theory. Such a specification would amount to determining a *particular state* of the system under scrutiny. The whole set of these states can be rearranged to yield a convex set. The path we have followed here is illustrated in Fig. 2, with the classical and quantum cases as extreme examples of a vast family.

The method described in this work can be seen as a general epistemological background for a huge family of scientific theories. In order to have a (quantitative) scientific theory, we must be able to make predictions on certain events of interest. Events are regarded here as propositions susceptible to being tested. Thus, one starts from an inferential calculus which allows for quantifying the degree of belief on the certainty of an event x, if it is known that event y has occurred. The crucial point is that event structures are not always organized as Boolean lattices (QM being perhaps the most spectacular example). Thus, in order to determine the general properties of the probabilities of a given theory (and thus all possible states by specifying particular values of prior probabilities on the atoms), we must apply Cox's method to lattices more general than Boolean ones.

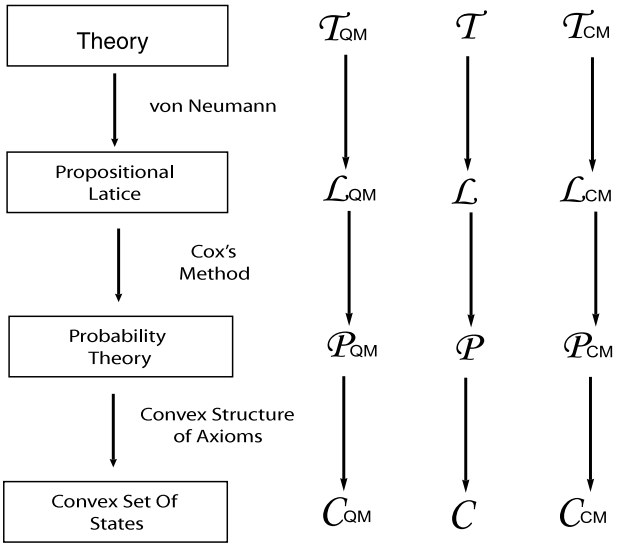

Fig. 2. Schematic representation of the method proposed here. A general theory $\mathcal{T}$ determines via the von Neumann approach the algebraic structure of the set of elementary tests. Then, by applying Cox's method, it is possible to determine the general properties of the canonical probability theory assigned to $\mathcal{T}$. Next, by assigning particular values to prior probabilities of atoms, all states will be determined and thus, the form of $\mathcal{C}$, the convex set of states. The quantum mechanical (QM) and the classical (CM) case are shown as the extreme instances of a vast family of theories.

7. Conclusions

By complementing the results presented in [23,20–22,24,15,26], in this work we showed that it is possible to combine Cox's method for the foundations of classical probability theory and the OQL approach, in order *to give an alternative derivation of what is known as non-Kolmogorovian (or non commutative) probability theory* [4,8].

Most physical probabilistic theories of interest can be endowed with orthomodular lattices. The elements of these lattices represent the events of that theory. As in [21,22], we have studied how the algebraic properties of the lattices determine – up to rescaling – the general form of the probabilities associated to the given physical theory. In this way, we provided a new formulation of the approach to physics based on non-Kolmogorovian probability theory [4].

Differently from [23,20–22,24,15,26], in this work we have focused on the application of the Cox's method to the von Neumann's lattice of projection operators in a separable Hilbert space and general orthomodular lattices as well, studying their particularities. In doing so, we obtained the von Neumann's axioms of quantum probabilities, and thus (using Gleason's theorem), quantum probabilities. In this way, our approach includes quantum mixtures explicitly and naturally.

It is interesting to remark (as in [21,22]) that this construction is not necessarily restricted to physics; any probabilistic theory endowed with a lattice structure of events will follow the same route.

We have also found that the method can be easily extended to σ-orthomodular posets (which are not lattices in the general case), and thus, it contains generalizations of QM.

In deriving quantum probability out of the lattice properties, we have shown in a direct way how non-distributivity forbids the derivation of a Kolmogorovian probability. This sheds light on the structure of quantum probabilities and on their differences with the classical case. In particular, we have shown that *the properties of the underlying algebraic structure imply that the inclusion–exclusion principle is not valid for the von Neumann's lattice*. And, moreover, that it can also fail in the more general framework of orthomodular (non-distributive) lattices. We have shown the explicit violation of the inclusion–exclusion principle for the finite example of the Chinese lantern and also for a non trivial relationship between atoms.

Furthermore, we have provided a non-trivial connection between Cox's approach and the problem posed by von Neumann regarding the axiomatization of probabilities. This is a novel perspective on the origin and axiomatization of the probabilities which appear in quantum theory. We believe that – far from being a definitive answer to the interpretation of quantum probabilities – this is an interesting topic to be further investigated.

Using Cox's approach, Shanon's entropy can be deduced as a natural measure of information over the Boolean algebra of classical propositions [18,26,24]. We will provide a detailed study of what happens with orthomodular latices elsewhere.

The strategy followed in this work suggests that we are at the gates of a great generalization. The general rule for constructing probabilities would read as follows:

- 1—We start by identifying the operational logic of our physical system. The characteristics of this "empirical" logic depends both on physical properties of the system and on the election of the properties that we assume in order to study the system. This can be done in a standard way, and the method is provided by the OQL approach.
- 2—Once the operational logic is identified, the symmetries of the lattice are used to define the properties of the "degree of implication" function, which will turn out to be the probability function associated to that particular logic. Remark that the same physical system may have different propositional structures, depending of the election of the observers. For example, if we look at the observable "electron's" charge, we will face classical propositions, but if wee look at its momentum and position, we will have a non-Boolean lattice.

This method is not only of physical interest, but also of mathematical interest, because one is solving the problem of characterizing probability measures over general lattices (and other structures as well). A final remark: our approach deviates from that of Cox, in the sense that we look for an empirical logic which would be intrinsic to the system under study, and because of that, not only referred to our ignorance about it, but to assumptions about its nature. However, this is not a problem, but an advantage: any problem which can be transformed into the language of lattice theory (or the more general framework of σ-orthocomplemented orthomodular posets) may fall into this scheme, and a probability theory can be developed using the method described in this work.

Acknowledgments

This work was partially supported by the grants PIP No 6461/05 and 1177 (CONICET). Also by the projects FIS2008-00781/FIS (MICINN) — FEDER (EU) (Spain, EU). The contribution of F. Holik to this work was done during his postdoctoral stance (CONICET) at Universidad Nacional de La Plata, Instituto de Física (IFLP-CCT-CONICET), C.C. 727, 1900 La Plata, Argentina.

References

[1] G. Birkhoff, J. von Neumann, Ann. Math. 37 (1936) 823–843.
[2] Dov M. Gabbay, John Woods, The Many Valued and Nonmonotonic Turn in Logic, Elsevier, Amsterdam, 2007, p. 205.
[3] M. Redei, in: K. Engesser, D.M. Gabbay, D. Lehmann (Eds.), The Birkhoff–von Neumann Concept of Quantum Logic, in: Handbook of Quantum Logic and Quantum Structures, Elsevier, 2009.
[4] S.P. Gudder, Stochastic Methods in Quantum Mechanics, North Holland, New York–Oxford, 1979.
[5] S.P. Gudder, in: A.R. Marlow (Ed.), Mathematical Foundations of Quantum Theory, Academic, New York, 1978.
[6] M.L. Dalla Chiara, R. Giuntini, R. Greechie, Reasoning in Quantum Theory, Kluwer Acad. Pub, Dordrecht, 2004.
[7] M. Rédei, Quantum Logic in Algebraic Approach, Kluwer Academic Publishers, Dordrecht, 1998.
[8] M. Rédei, S. Summers, Stud. Hist. Philos. Sci. B Stud. Hist. Philos. Modern Phys. 38 (2) (2007) 390–417.
[9] G. Mackey, Mathematical Foundations of Quantum Mechanics, W.A. Benjamin, New York, 1963.
[10] E. Davies, J. Lewis, Comm. Math. Phys. 17 (1970) 239–260.
[11] M. Srinivas, J. Math. Phys. 16 (1975) 1672.
[12] J. Acacio de Barros, P. Suppes, 2000. arXiv:quant-ph/0001017v1.
[13] C. Anastopoulos, Ann. Physics 313 (2004) 368–382.
[14] J. Rau, Ann. Physics 324 (2009) 2622–2637.
[15] P. Goyal, K. Knuth, Symmetry 3 (2) (2011) 171–206. http://dx.doi.org/10.3390/sym3020171.
[16] A.N. Kolmogorov, Foundations of Probability Theory, Julius Springer, Berlin, Germany, 1933.
[17] R.T. Cox, Amer. J. Phys. 14 (1946) 1–13.
[18] R.T. Cox, The Algebra of Probable Inference, The Johns Hopkins Press, Baltimore, MD, USA, 1961.

[19] G. Boole, An Investigation of the Laws of Thought, Macmillan, London, UK, 1854.
[20] K.H. Knuth, in: G.J. Erickson, Y. Zhai (Eds.), Proceedings of 23rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, in: Bayesian Inference and Maximum Entropy Methods in Science and Engineering, American Institute of Physics, New York, NY, USA, 2004, pp. 204–235.
[21] K.H. Knuth, in: P. Goggans, C.Y. Chan (Eds.), Proceedings of 23rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Vol. 707, in: Bayesian Inference and Maximum Entropy Methods in Science and Engineering, American Institute of Physics, New York, NY, USA, 2004, pp. 132–144.
[22] K.H. Knuth, Valuations on lattices and their application to information theory, in: Proceedings of the 2006 IEEE World Congress on Computational Intelligence, Vancouver, Canada, 2006. July.
[23] A. Caticha, Phys. Rev. A 57 (1998) 1572–1582.
[24] K.H. Knuth, Neurocomputing 67C (2005) 245–274.
[25] P. Goyal, K.H. Knuth, J. Skilling, Phys. Rev. A 81 (2010) 022109.
[26] K.H. Knuth, Skilling, 2012. arXiv:1008.4831v1.
[27] G.W. Mackey, Amer. Math. Monthly Supplement 64 (1957) 45–57.
[28] J.M. Jauch, Foundations of Quantum Mechanics, Addison-Wesley, Cambridge, 1968.
[29] C. Piron, Foundations of Quantum Physics, Addison-Wesley, Cambridge, 1976.
[30] G. Kalmbach, Orthomodular Lattices, Academic Press, San Diego, 1983.
[31] G. Kalmbach, Measures and Hilbert Lattices, World Scientific, Singapore, 1986.
[32] V. Varadarajan, Geometry of Quantum Theory I, van Nostrand, Princeton, 1968.
[33] V. Varadarajan, Geometry of Quantum Theory II, van Nostrand, Princeton, 1970.
[34] J.R. Greechie, in: E. Beltrameti, B. van Fraassen (Eds.), Current Issues in Quantum Logic, Plenum, New York, 1981, pp. 375–380.
[35] R. Giuntini, Quantum Logic and Hidden Variables, BI Wissenschaftsverlag, Mannheim, 1991.
[36] P. Pták, S. Pulmannova, Orthomodular Structures as Quantum Logics, Kluwer Academic Publishers, Dordrecht, 1991.
[37] E.G. Beltrametti, G. Cassinelli, The Logic of Quantum Mechanics, Addison-Wesley, Reading, 1981.
[38] A. Dvurečenskij, S. Pulmannová, New Trends in Quantum Structures, Kluwer Acad. Pub., Dordrecht, 2000.
[39] K. Engesser, D.M. Gabbay, D. Lehmann (Eds.), Handbook of Quantum Logic And Quantum Structures (Quantum Logic), North-Holland, 2009.
[40] D. Aerts, I. Daubechies, Lett. Math. Phys. 3 (1979) 11–17.
[41] D. Aerts, I. Daubechies, Lett. Math. Phys. 3 (1979) 19–27.
[42] D. Aerts, J. Math. Phys. 24 (1983) 2441.
[43] D. Aerts, Rep. Math. Phys. 20 (1984) 421–428.
[44] D. Aerts, J. Math. Phys. 25 (1984) 1434–1441.
[45] B. Mielnik, Comm. Math. Phys. 9 (1968) 55–80.
[46] B. Mielnik, Comm. Math. Phys. 15 (1969) 1–46.
[47] B. Mielnik, Comm. Math. Phys. 37 (1974) 221–256.
[48] G. Ludwig, Comm. Math. Phys. 4 (1967) 331–348.
[49] G. Ludwig, Comm. Math. Phys. 9 (1968) 1–12.
[50] J. von Neumann, Mathematical Foundations of Quantum Mechanics, twelth ed., Princeton University Press, Princeton, 1996.
[51] M. Reed, B. Simon, Methods of modern mathematical physics I: Functional analysis, Academic Press, New York-San Francisco-London, 1972.
[52] M. Solèr, Comm. Algebra 23 (1995) 219–243.
[53] A. Wilce, in: Edward N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Spring 2009 Edition), in: Quantum Logic and Probability Theory, Archive edition: Spring, 2009. URL = http://plato.stanford.edu/archives/spr2009/entries/qt-quantlog/.
[54] H. Barnum, J. Barret, M. Leifer, A. Wilce, Phys. Rev. Lett. 99 (2007) 240501.
[55] H. Barnum, A. Wilce, arXiv:0908.2352v1 [quant-ph], 2009;
Electron. Notes Theor. Comput. Sci. 270 (1) (2011) 3–15.
[56] H. Barnum, R. Duncan, A. Wilce, arXiv:1004.2920v1 [quant-ph], 2010.
[57] M. Rédei, Math. Intelligencer 21 (4) (1999) 7–12.
[58] J. Aczél, Lectures on Functional Equations and Their Applications, Academic Press, New York, 1966.
[59] A. Gleason, J. Math. Mech. 6 (1957) 885–893.
[60] D. Buhagiar, E. Chetcuti, A. Dvurečenskij, Found. Phys. 39 (2009) 550–558.
[61] N. Chomsky, IRE Trans. Inform. Theory 2 (1956) 113.
[62] K. Svozil, Quantum Logic, Springer-Verlag, Singapore, 1998.