Review

Generalized Probabilities in Statistical Theories

Federico Holik [1,*], **César Massri** [2] , **Angelo Plastino** [1] and **Manuel Sáenz** [3]

1 Instituto de Física (IFLP-CCT-CONICET), Universidad Nacional de La Plata, C.C. 727,
La Plata 1900, Argentina; angeloplastino@gmail.com
2 Departamento de Matemática, Universidad CAECE, Buenos Aires 1084, Argentina; cmassri@caece.edu.ar
3 International Center for Theoretical Physics, UNESCO, 34151 Trieste, Italy; saenz.manuel@gmail.com
* Correspondence: holik@fisica.unlp.edu.ar

Abstract: We discuss different formal frameworks for the description of generalized probabilities in statistical theories. We analyze the particular cases of probabilities appearing in classical and quantum mechanics and the approach to generalized probabilities based on convex sets. We argue for considering quantum probabilities as the natural probabilistic assignments for rational agents dealing with contextual probabilistic models. In this way, the formal structure of quantum probabilities as a non-Boolean probabilistic calculus is endowed with a natural interpretation.

1. Introduction

In the year 1900, the great mathematician David Hilbert presented a famous list of problems at a Conference in Paris. Hilbert suggested that the efforts of the mathematicians in the years to come should be oriented in the solution of these problems. The complete list was published later [1]. Remarkably, one of these problems was dedicated to the axiomatic treatment of probability theory and physical theories. In Hilbert's own words ([1], p. 454):

> "The investigations on the foundations of geometry suggest the problem: To treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important part; in the first rank are the theory of probabilities and mechanics. As to the axioms of the theory of probabilities, it seems to me desirable that their logical investigation should be accompanied by a rigorous and satisfactory development of the method of mean values in mathematical physics, and in particular in the kinetic theory of gases."

After a series of preliminary investigations by many researchers (see, for example [2]), an axiomatization of probability theory was finally presented in the 1930s by Andrey Kolmogorov [3]. This contribution, which can be considered as the foundation of modern probability theory, is based on measure theory. Indeed, in Kolmogorov's axiomatic treatment, probability is considered as a measure defined over a suitable collection of events, organized as a *sigma-algebra* (which is found to be also a *Boolean lattice*). His list of axioms allows the description of many examples of interest and was considered as a reasonable fulfillment of Hilbert's program for probability theory.

Hilbert himself dedicated great efforts to solve his sixth problem. His contributions were influential in the development of Relativity Theory, and he also contributed to the development of Quantum Mechanics. Indeed, Quantum Mechanics acquired its rigorous axiomatic formulation after a series of papers by Hilbert, J. von Neumann, L. Nordheim, and E. P. Wigner [4]. It can be said that its definitive form was accomplished in the book of von Neumann [5]. This axiomatic approach was extended to the relativistic setting in subsequent years (see, for example, [6,7]; see [8] for a more updated exposition of

the algebraic approach; and for a rigorous formulation of quantum statistical mechanics, see [9]).

However, the advent of Quantum Mechanics presented a model of probabilities that had many peculiar features. R. P. Feynman stated this clearly in [10], p. 533:

> "I should say, that in spite of the implication of the title of this talk the concept of probability is not altered in quantum mechanics. When I say the probability of a certain outcome of an experiment is *p*, I mean the conventional thing, that is, if the experiment is repeated many times one expects that the fraction of those which give the outcome in question is roughly *p*. I will not be at all concerned with analyzing or defining this concept in more detail, for no departure of the concept used in classical statistics is required. What is changed, and changed radically, is the method of calculating probabilities."

What is the meaning of Feynman's words? Feynman tells us that the way of computing frequencies is not altered in quantum mechanics: the real numbers yielded by Born's rule can be tested in the lab in the usual way. However, the method for *computing* probabilities has changed in a radical way. As put in [11], this can be rephrased as follows: the radical change has to do with the recipe that quantum mechanics gives us for calculating new probabilities from old. The radical change mentioned by Feynman lies behind all the astonishing features of quantum phenomena. This was recognized very quickly as a non-classical feature. These peculiarities and the formal aspects of the probabilities involved in quantum theory have been extensively studied in the literature [12–19]. We refer to the probabilities related to quantum phenomena as *quantum probabilities* (QP). Accordingly, we refer to probabilities obeying Kolmogorov's axioms as *classical probabilities* (CP).

In this paper, we discuss the formal structure of quantum probabilities as measures over a non-Boolean algebra. We focus on a crucial aspect of quantum probabilities—namely, that there exists a major structural difference between classical states and quantum states:

- States of classical probabilistic systems can be suitably described by Kolmogorovian measures. This is due to the fact that each classical state defines a measure in the Boolean sigma-algebra of measurable subsets of phase space.
- Contrarily to classical states, quantum states cannot be reduced to a single Kolmogorovian measure. A density operator representing a quantum state defines a measure over an orthomodular lattice of projection operators, which contains (infinitely many) incompatible maximal Boolean subalgebras. These represent different and complementary—in the Bohrian sense—experimental setups. The best we can do is to consider a quantum state as a family of Kolmogorovian measures, pasted in a harmonic way [20]; however, there is no joint (classical) probability distribution encompassing all possible contexts.

We discuss the above mentioned differences in relation to quantum theory as a non-Kolmogorovian probability calculus. This calculus can be considered as an extension of classical measure theory to a non-commutative setting (see, for example [12,21]; see also [22] for a study of quantum measure theory). In this way, the axiomatization of probabilities arising in QM (and more general probabilistic models) can be viewed as a continuation of the Hilbert's program with regard to probability theory. We argue that the probabilities in generalized probabilistic models can be interpreted, in a natural way, in terms of reasonable expectations of a rational agent facing event structures that may define different and incompatible contexts. This allows us to understand other related notions, such as random variables and information measures, as natural generalizations of the usual ones.

Kolmogorov's approach to probability theory is not the only one. In the attempts to establish foundations for probability, we have to mention the works of de Finetti [23] and R. T. Cox [24,25] (in connection with R. T. Cox works, see also [26]). For a detailed and accessible study of the history of probability theory and its interpretations, we refer the reader to the Apendix of [2]. In this paper, we pay special attention to Cox's approach and

make use of its extension to the quantum realm [27]. Cox's approach is based on a study of the measure functions compatible with the algebraic properties of the logic of a rational agent trying to make probable inferences out of the available data. Different variants of this approach have been used to describe probabilities in QM [27–34].

In [27], it is shown that the peculiar features of QP arise whenever the lattice of propositions of Cox's approach is replaced by a non-distributive one. As is well known, the standard quantum-logical approach to QM characterizes this theory using a lattice-theoretical framework in which the lattices are orthomodular [21,35–53]. In [27], and it is shown that, when Cox's method is applied to the orthomodular lattice of projections of a Hilbert space, QP are derived.

We remark that generalized probabilities can also be studied in what has been called the Convex Operational Models (COM) approach [54–64]. These are called "generalized probabilistic theories" (GPTs) or simply, "operational models" (see also [17]). Different mathematical frameworks are used to describe GPTs. Here, we focus on the most important ones, paying attention to how they can be inter-translated (when possible). In the COM 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, and many information protocols) can be suitably described using this approach [55–64]. Non-linear generalizations of QM were studied using the convex approach in [65–67].

It is important to understand the relations between the different formulations of GPTs. For example, the measures over complete orthomodular lattices discussed in Section 4 of this work define GPTs (while an arbitrary GPT might be not be describable in terms of a measure over an orthomodular lattice). The reason why models defined over lattices are so important is that all relevant physical theories can be described in such a setting. Indeed, all relevant physical models can be ultimately described using von Neumann algebras, that are generated by their lattices of projection operators (see, for example [8,9,21]). This is the case for classical statistical theories, standard quantum mechanics, relativistic quantum mechanics, and quantum statistical mechanics. As an example (as we will discuss in Section 3.2), states of models of quantum mechanics can be described as measures over the orthomodular lattice of projection operators acting on a separable Hilbert space. It is interesting, for several reasons, to study more general models (that could describe, for example, alternative physical theories). However, there is always a trade off between generality and particularities: if our models are too general, we can loose valuable information about the geometric and algebraic structures involved in relevant physical theories. On the contrary, if they are too specific, we might loose information about the general road map for exploration. It is our aim here to shed some light onto this vast field of research, putting the focus on the idea that Kolmogorov's framework can certainly be generalized in a reasonable and useful way.

It is also important to mention that quantum-like probabilities have been considered outside the quantum domain. Many probabilistic systems behave in a contextual way, and then, it is reasonable to attempt to use quantum-like probabilities to described them, since these are specially suited to deal with contextual behavior. This is an exciting field of research that has grown intensively during recent years (see, for example [68–70]).

We start by reviewing different approaches to CP, namely, Kolmogorov's and Cox's, in Section 2. Next, we discuss the formalism of QM in Section 3, emphasizing how it can be considered as a non-Boolean version of Kolmogorov's theory. In Sections 4 and 5, we discuss generalizations using orthomodular lattices and COMs, respectively. After discussing alternative approaches in Section 6, we present the generalization of the Cox method to general non-distributive lattices in Section 7. Finally, our conclusions are drawn in Section 8. Given that lattice theory is so central to the discussions presented here, we have included a short review about its elementary notions in Appendix A.

2. Classical Probabilities

This Section is devoted to *classical probability theory* (CP). However, 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; however, we will discuss two of the most important approaches to CP. These are the one given by A. N. Kolmogorov [3] and the one given by R. T. Cox [24,25].

2.1. Kolmogorov

Kolmogorov presented his axiomatization of classical probability theory [3] in the 1930s. It can be formulated 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

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

satisfying

$$\mu(\Omega) = 1, \tag{1b}$$

and, if I is a denumerable set of indices, for any pairwise disjoint family $\{A_i\}_{i \in I}$,

$$\mu(\bigcup_{i \in I} A_i) = \sum_i \mu(A_i) \tag{1c}$$

Conditions (1a)–(1c) are known as *axioms of Kolmogorov* [3]. The triad (Ω, Σ, μ) is called a *probability space*. Probability spaces obeying Equations (1a)–(1c) are usually referred as Kolmogorovian, classical, commutative, or Boolean probabilities [16], due to the Boolean character of the σ-algebra in which they are defined.

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

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \tag{2}$$

or, as expressed in logical terms, by replacing "$\vee$" instead of "$\cup$" and "$\wedge$" instead of "$\wedge$":

$$\mu(A \vee B) = \mu(A) + \mu(B) - \mu(A \wedge B) \tag{3}$$

As remarked in [71], Equation (2) 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 for each event to occur in a series of N repetitions, then (2) trivially holds. Notice that (3) implies that:

$$\mu(A) + \mu(B) \leq \mu(A \vee B) \tag{4}$$

The inequality (4) no longer holds in QM, a fact linked to its non-Boolean character (see, for example [16]), Section 2.2. Indeed, for a suitably chosen state and events A and B (i.e., for a non-commutative pair), (4) can be violated. If $N(A \vee B)$, $N(A)$, $N(B)$ and $N(A \wedge B)$ are the number of times for each event to occur in a series of N repetitions, then the sum rule should trivially hold (but it does not). This poses problems to a relative-frequencies' interpretation of quantum probabilities (see, for example, the discussion posed in [71]). The QM example shows that non-distributive propositional structures give rise to probability theories that appear to be very *different from* those of Kolmogorov. Notwithstanding, it is important to mention that some authors have managed to develop a relative frequencies interpretations (see, for example [72]).

If all possible measures satisfying (1a)–(1c) are considered as forming a set $\Delta(\Sigma)$ (with Σ fixed), then it is straightforward to show that it is convex. As we shall see below, it is a simplex, and its form will be related to the Boolean character of the lattice of classical events.

2.2. Random Variables and Classical States

It is important to recall here how random variables are defined in Kolmogorov's setting (according to the measure theoretic approach). See [16] for a detailed exposition. A random variable f can be defined as a measurable function $f : \Omega \longrightarrow \mathbb{R}$. In this context, by a measurable function f, we mean a function satisfying that, for every Borel subset B of the real line (the *borel sets* ($B(\mathbb{R})$) are defined as the family of subsets of $\mathbb{R}$ such that (a) it is closed under set theoretical complements, (b) it is closed under denumerable unions, and (c) it includes all open intervals [73]), we have that $f^{-1}(B) \in \Sigma$ (i.e., the pre-image of every Borel set B under f belongs to Σ, and thus it has a definite probability measure given by $\mu(f^{-1}(B))$).

Notice that a random variable f defines an inverse map f^{-1}

$$f^{-1} : B(\mathbb{R}) \longrightarrow \Sigma \tag{5a}$$

satisfying

$$f^{-1}(\varnothing) = \varnothing \tag{5b}$$

$$f^{-1}(\mathbb{R}) = \Omega \tag{5c}$$

$$f^{-1}(\bigcup_j B_j) = \bigcup_j f^{-1}(B_j) \tag{5d}$$

for any disjoint denumerable family of Borel sets B_j. Denoting the complement of a set X by X^c, we have that, for every Borel set B:

$$f^{-1}(B^c) = (f^{-1}(B))^c. \tag{5e}$$

To illustrate ideas, let us consider a classical probabilistic system. A classical observable H (such as the energy) will be a function from the state space Γ to the real numbers. The state of the system, given by a probability density ϱ (i.e., $\varrho \geq 0$ and with Lebesgue integral $\int_\Gamma \varrho(x)dx = 1$)), will define a measure μ over the measurable subsets of Γ as follows. For each subset $S \in \Gamma$, define

$$\mu(S) := \int_S \varrho(x)dx. \tag{6}$$

Measurable subsets of Γ will be those for which the above integral converges. The function μ will obey Kolmogorov's axioms, provided that we take $\Gamma = \Omega$ and Σ as the set of measurable subsets of Γ. The above formula is sufficient to compute the mean values and the probabilities of any event of interest. Given an elementary testable proposition such as: "the values of the observable H lie in the interval (a, b)", the real number $\mu(H^{-1}((a, b)))$ gives us the probability that this proposition is true. In this sense, each observable of a classical probabilistic system can be considered as a random variable. This has to be a necessary condition for any admissible classical state: a state must specify definite probabilities for every elementary test that we may perform on the system. In this sense, each classical (probabilistic) state can be described by a Kolmogorovian measure, with the observables represented as random variables.

At the same time, by associating "$\vee$" with "$\cup$", "$\wedge$" with "$\cap$", "$\neg$" with "$(\ldots)^c$" (set theoretical complement), and "$\leq$" with "$\subseteq$" (set theoretical inclusion), we see that the Boolean structure associated to measurable subsets is coincident with the distributive character of classical logic. The fact that the logic associated to a classical system is Boolean (in the above operational sense), was one of the main observations in [35].

As we will see in the following sections, the quantum formalism can be considered as an extension of the classical one, provided that we replace the measurable subsets of phase space with $\mathcal{P}(\mathcal{H})$ (the lattice of projection operators on a Hilbert space $\mathcal{H}$), the measure μ by a quantum state represented by a density operator, and the classical random variables by projection valued measures associated to self-adjoint operators. As a consequence, the operational logic associated to a quantum system will fail to be Boolean [35], due to the

non-distributive character of $\mathcal{P}(\mathcal{H})$. The set of states of a quantum system will be convex too. However, the geometrical shape will be very different from that of a classical one, due to the non-Boolean character of the lattice of events involved.

2.3. Cox's Approach

Since the beginning of probability theory, there has been a school of thought known as Bayesianism, which treats probabilities in a different manner from the one discussed in the previous section. For the Bayesian approach, probabilities are 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 probability theory [74]. In his work, Laplace proposed a way to assign probabilities in situations of ignorance that would eventually be known as the"Laplace principle". Later works would attempt to formalize and give coherence to the Bayesian approach, as for example, [23,75]. In this section, we center our attention on one of these attempts [24,25], of R. T. Cox.

While attaining equivalent results to those of Kolmogorov, Cox's approach is conceptually very different. In the Kolmogorovian approach, probabilities can be naturally interpreted (though 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 a rational agent—which may 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 for inferential reasoning that would be coherent with classical logic and that would reduce to it whenever all the premises have definite truth values.

To do this, he started with two very general axioms and presupposed the calculus of classical propositions, which, as is well known, forms a Boolean lattice [76]. By doing so, he derived classical probability theory as an inferential calculus on Boolean lattices. We sketch here the arguments presented in his book [24]. For a more detailed exposition on the deductions, the reader is referred to [24,25,30,31,33,77,78]. See [79] for discussions on a rigorization of Cox's method.

The two axioms used by Cox [24] are

- C1—The probability of an inference on given evidence determines the probability of its contradiction from the same evidence.
- C2—The probability on a given evidence that both of two inferences are true is determined by their separate probabilities, one from the given evidence and the other from 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 regarding how likely it is that a is true given that the agent knows that the hypothesis h is true.

Then, requiring the function φ to be coherent with the properties of the calculus of classical propositions, the agent 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 $F[\varphi(a|h), \varphi(b|h)] \equiv \varphi(a \wedge b|h) : \mathbb{R}^2 \to \mathbb{R}$, the agent arrives at a functional equation for $F(x, y)$:

$$F[x, F(y, z)] = F[F(x, y), z] \tag{7}$$

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:

$$P(a \wedge b|h) = CP(a|h \wedge b)P(b|h) \tag{8}$$

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 $f[P(a|h)] \equiv P(\neg a|h) : \mathbb{R} \to \mathbb{R}$, we arrive at the following functional equation for $P(a|h)$

$$P(a|h)^r + P(\neg a|h)^r = 1 \qquad (9)$$

With r as 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 decided to take $r = 1$ and to continue using $P(a|h)$.

Using Equations (7) and (9), the law of double negation and Morgan's law for conjunction ($\neg(a \wedge b) = \neg a \vee \neg b$), we arrive at the *sum rule* of probability theory:

$$P(a \vee b|h) = P(a|h) + P(b|h) - P(a \wedge b|h) \qquad (10)$$

As it turns out, $P(a|h)$—if suitably normalized—satisfies the properties of an additive Kolmogorovian probability (Equations (1a)–(1c)).

Due to the importance of Cox's theorem to the foundations of probability, it has been the target of 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 (7). Since then, there have been different proposals to save Cox's approach by proving it using less restrictive axioms. In [80], 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 [78].

Once the general properties of the function $P(a|h)$ are established, the next problem is to find a way to determine prior probabilities (i.e., probabilities conditional only to the hypothesis h). Although, formally, one could assign prior probabilities in any way coherent with the normalization used, 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 [26,77], which we will review shortly in the next section. Other ways of assigning prior probabilities include the Laplace principle [75] and coherence with symmetry transformations [81]. Nevertheless, the existence of a general algorithm for assigning prior probabilities is still an open question.

2.4. 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 assuming more information than that 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, the 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 as the uniform distribution, and the MaxEnt principle reduces to Laplace's. Different kinds of constraints result in different prior probability distributions (PPD) [26]. In [82], a table of some of the distributions obtained in this way is presented. Although, given a set of constraints, the corresponding PPD can be readily computed, there is no general method 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 [77]. Assuming that the prior knowledge about the system is given by n expectation values of a collection 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 maximizing Shannon's logarithmic entropy S subject to the n constraints

$$\langle R_i \rangle = r_i; \text{ for all } i. \tag{11}$$

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 [77]. $S(maximal)$, once determined, yields complete thermodynamical information with respect to the system of interest [77]. The MaxEnt probability distribution function (PDF), associated to Boltzmann–Gibbs–Shannon's logarithmic entropy S, is given by [77]

$$\rho_{max} = \exp\left(-\lambda_0 \mathbf{1} - \lambda_1 R_1 - \cdots - \lambda_n R_n\right), \tag{12}$$

where the λ's are Lagrange multipliers guaranteeing that

$$r_i = -\frac{\partial}{\partial \lambda_i} \ln Z, \tag{13}$$

while the partition function reads

$$Z(\lambda_1 \cdots \lambda_n) = \sum_i \exp^{-\lambda_1 R_1(x_i) - \cdots - \lambda_n R_n(x_i)}, \tag{14}$$

and the normalization condition

$$\lambda_0 = \ln Z. \tag{15}$$

In a quantum setting, the R's are operators on a Hilbert space $\mathcal{H}$, while ρ is a density matrix (operator). The sum in the partition function must be replaced by a trace, and Shannon's entropy must be replaced by von Neumann's.

3. The Formalism of QM

In this Section, we discuss some specific features of the quantum formalism [5,21,38,39] that are relevant for the problem of QP.

3.1. Elementary Measurements and Projection Operators

In QM, observable physical magnitudes are represented by compact self-adjoint operators in a Hilbert space $\mathcal{H}$ (we denote this set by $\mathcal{A}$). Due to the spectral decomposition theorem [5,73], a key role is played by the notion of *projection valued measure* (PVM): the set of PVMs 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 that assigns a projection operator to each interval of the real line. In this sense, projection operators are the building blocks out of which any observable can be built. It is important to recall that projection operators have a very clear operational meaning: they represent elementary empirical tests with only two outputs (zero and one, or YES and NO). In a formal way, a PVM is a map M defined over the Borel sets (see Section 2.2) as follows

$$M : B(\mathbb{R}) \to \mathcal{P}(\mathcal{H}), \tag{16a}$$

satisfying

$$M(\varnothing) = \mathbf{0} \ \ (\mathbf{0} := \text{null subspace}) \tag{16b}$$

$$M(\mathbb{R}) = \mathbf{1} \tag{16c}$$

$$M(\cup_j (B_j)) = \sum_j M(B_j), \tag{16d}$$

for any disjoint denumerable family B_j.

$$M(B^c) = \mathbf{1} - M(B) = (M(B))^\perp \tag{16e}$$

As we will see in the following Section, a PVM is the natural generalization of the notion of random variable to the non-Boolean setting. In order to realize why this is so, it is important to compare Equations (5a)–(5e) and (16a)–(16e). It is also important to remark that the set of projections in the image of a PVM are always orthogonal: this implies that this set can always be endowed with a Boolean lattice structure. This allows us to associate, to each complete observable, a particular empirical context represented by a Boolean algebra of events. Thus, in this sense, complete observables are always referred to a particular context.

Fixing an element $A \in \mathcal{A}$, the intended interpretation of the associated PVM ($M_A(\ldots)$), evaluated in an interval (a, b) (i.e., $M_A((a, b)) = P_{(a,b)}$) is: "the value of A lies between the interval (a, b)". In this sense, projection operators represent elementary tests or propositions in QM. In other words, they can be considered as the simplest quantum mechanical observables. As we reviewed in Appendix A, 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 [35]. We refer to it as the *von Neumann-lattice* ($\mathcal{P}(\mathcal{H})$) [21]. As shown in [35], an analogous treatment can be done for classical systems. As we have seen in Section 2.2, propositions associated to a classical system are endowed with a natural Boolean structure.

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* [21,83–86]), as an attempt of generalizing certain algebraic properties of Jordan algebras [4]. 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 [83–86]. On the other hand, lattice theory is deeply connected to projective geometry [87], 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) [12,21].

The quantum logical approach of Birkhoff and von Neumann was continued by other researchers [16,17,38,39,42,43,50] (see [36,41,44] for complete expositions). One of the key results of this approach is the *representation theorem* of C. Piron [43]. He showed that any propositional system can be coordinatized in a generalized Hilbert space. A later result by Solèr showed that, by adding extra assumptions, it can only be a Hilbert space over the fields of the real numbers, complex numbers, or quaternions [88].

3.2. Quantum States and Quantum Probabilities

In this Section we discuss QP. We do this by reviewing the usual approach, in which Kolmogorov's axioms are extended to non-Boolean lattices (or algebras) [12].

As we have seen in Section 3.1, elementary tests in QM are represented by closed subspaces of a 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 that are similar to those of Kolmogorov [5,21,37]. The Boolean Σ-algebra appearing in Kolmogorov's axioms (Equations (1a)–(1c)) is replaced by $\mathcal{P}(\mathcal{H})$, and a measure s is defined as follows:

$$s : \mathcal{P}(\mathcal{H}) \to [0; 1] \tag{17a}$$

such that:

$$s(\mathbf{1}) = 1 \ (\mathbf{1} := \mathcal{H}) \tag{17b}$$

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

$$s\left(\sum_j P_j\right) = \sum_j s(P_j). \tag{17c}$$

In this way, a real number between 0 and 1 is assigned to any elementary test. Despite the similarity with Kolmogorov's axioms, the probabilities defined above are very different, due to the non-Boolean character of the lattice involved. Gleason's theorem [89,90] asserts that if $dim(\mathcal{H}) \geq 3$, any measure s satisfying (17a)–(17c) can be put in correspondence with a trace class operator (of trace one) ρ_s:

$$s(P) := \mathrm{tr}(\rho_s P) \tag{18}$$

for every orthogonal projection P. On the other hand, using Equation (18), any trace class operator of trace one defines a measure as in (17a)–(17c), and thus the correspondence is one to one for $dim(\mathcal{H}) \geq 3$ (something that is not true for the two dimensional case). In this way, Equations (17a)–(17c) define the usual probabilities of QM and constitute a natural generalization of Kolmogorov's axioms to the quantum case.

The set $\mathcal{C}(\mathcal{H})$ of all possible measures satisfying Equations (17a)–(17c) is indeed convex, as in the classical case. However, these sets are very different. As an example, let us compare a classical bit (to fix ideas, think about the possible probabilistic states of a coin) and a qubit (a quantum system represented by a two-dimensional model). While the state space of the first one is a line segment, it is well known that the state space of the second is homeomorphic to a sphere [91]. For more discussion about the convex set of quantum states, see [92,93].

A state satisfying Equations (17a)–(17c) will yield a Kolmogorovian probability when restricted to a maximal Boolean subalgebra of $\mathcal{P}(\mathcal{H})$. In this way, a quantum state can be considered as a coherent pasting of different Kolmogorovian measures. This has a natural physical interpretation as follows. Each empirical setup will define a maximal Boolean algebra. The fact that quantum states yield the correct observed frequencies (via the Born rule), allows defining consistent Kolmogorovian probabilities on each Boolean setting. However, doing statistics on repeated measurements in identical preparations using a single empirical setup, is not sufficient to determine a quantum state completely.

For a general state, it will be mandatory to perform measurements in different and complementary (in Bohr's sense) empirical setups. Notice that there are many ways in which one could define a family of Kolmogorovian probabilities in $\mathcal{P}(\mathcal{H})$. However, the probabilities defined by a quantum state (or equivalently, by Equations (17a)–(17c)), have a very particular mathematical form. The existence of uncertainty relations between non-compatible observables [94] is nothing but an expression of this fact. The fact that a quantum state μ can be considered as a coherent collection of Kolmogorovian measures can be summarized in a diagram as follows. If Σ is an arbitrary Boolean subalgebra of $\mathcal{P}(\mathcal{H})$, let μ be a quantum state and μ_Σ the restriction of μ to Σ. Then, for every Boolean subalgebra Σ, we have the following commutative diagram:

$$
\begin{array}{ccc}
\Sigma & \hookrightarrow & \mathcal{P}(\mathcal{H}) \\
& \searrow^{\mu_\Sigma} & \downarrow^{\mu} \\
& & [0,1]
\end{array}
$$
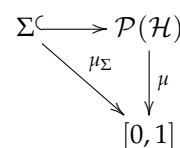
The fact that there exists a global quantum state μ that makes the above diagram commute for every Boolean subalgebra, is a quite remarkable fact about the quantum formalism. Notice that, given that the intersection of Boolean subalgebras may be non-trivial (see examples in the next Section), the probability assignments must satisfy certain compatibility

conditions. Thus, even if an event x belongs to two different measurement contexts, the quantum state assigns to it the same probability. That is, the probability is assigned independently of the context to which it belongs. This is known as the *no-signal* condition, which will not necessarily hold outside physics (for example, in cognition experiments).

The above generalization also includes quantum observables in a natural way. Indeed, by appealing to the spectral decomposition theorem, there is a one-to-one correspondence between quantum observables represented by self-adjoint operators and PVM's. Then, these notions are interchangeable. However, a quick look to Equations (16a)–(16e) reveals that PMS are very similar to classical random variables: while classical random variables map Borel sets into the Boolean lattice of measurable sets, PVMs map Borel sets into the non-Boolean lattice $\mathcal{P}(\mathcal{H})$. Thus, quantum observables can be reasonably interpreted as non-Kolmogorovian random variables.

We mentioned above that (1a)–(1c) and (17a)–(17c) are not equivalent probability theories. For example, Equation (2) is no longer valid in QM. Indeed, for suitably chosen s and quantum events A and B, we have

$$s(A) + s(B) \leq s(A \vee B) \tag{19}$$

The above inequality should be compared with the classical one, given by (4). As an example, consider a two dimensional quantum system (in current jargon: a qubit), the events $A = |\uparrow_z\rangle\langle\uparrow_z|$ (spin up in direction $\hat{z}$) and $B = |\downarrow_x\rangle\langle\downarrow_x|$ (spin down in direction $\hat{x}$), and the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_z + |\downarrow\rangle_z)$ ("cat state" in the basis $\hat{z}$, which is the same as "spin up in direction $\hat{x}$"). Thus, using some simple math, we obtain that $p_{|\psi\rangle}(A) = \mathrm{tr}(|\psi\rangle\langle\psi||\uparrow_z\rangle\langle\uparrow_z|) = \frac{1}{2}$, and $p_{|\psi\rangle}(B) = \mathrm{tr}(|\psi\rangle\langle\psi||\downarrow_x\rangle\langle\downarrow_x|) = 0$. On the other hand, we have that $A \vee B$ is the linear span of $|\uparrow_z\rangle$ and $|\downarrow_x\rangle$, and then, $A \vee B = I$. Since $p_{|\psi\rangle}(I) = \mathrm{tr}(|\psi\rangle\langle\psi|I) = 1$, we have that $p_{|\psi\rangle}(A) + p_{|\psi\rangle}(B) = \frac{1}{2} + 0 < 1 = p_{|\psi\rangle}(A \vee B)$.

The probability theory defined by (17a)–(17c) can also be considered as a non-commutative generalization of classical probability theory in the following sense: while when 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* [12,16]. We end this Section by noting that some technical complications appear when one attempts to define a quantum conditional probability in the non-commutative setting. For a complete discussion about these matters and a comparison between classical and quantum probabilities, see [12,16].

3.3. Some Examples

In order to understand better the mathematical structure (and the physical interpretation) underlying quantum probabilities, we discuss here some examples in detail. We relate the event structures associated to physical systems with different notions of lattice theory. We do this by enumerating different examples that are relevant for the discussion presented in this work. The reader unfamiliar with lattice theory can consult Appendix A.

Finite Probability model: a dice. Consider the throw of a dice. The possible outcomes are given by $\Omega = \{1, 2, 3, 4, 5, 6\}$. A probabilistic state of the dice is determined by assigning real numbers p_i, $i = 1, ..., 6$, to each element of Ω. If the dice is not loaded, then $p_i = \frac{1}{6}$ for all i; however, a realistic dice will not satisfy this. An event will be represented by a subset of Ω. As examples, consider the event "the outcome is even" or "the outcome is greater than 2". These are represented by $\{2, 4, 6\}$ and $\{3, 4, 5, 6\}$, respectively. All possible subsets of Ω form a Boolean lattice (see Apendix A), with regard to the set-theoretical operations: "$\cup$" (interpreted as "$\vee$"), "$\cap$" (interpreted as "$\wedge$"), and the set theoretical complement (interpreted as "$\neg$"). The example of a $\sigma-$algebra associated to a measurable space (Ω, Σ, μ) works in a similar way.

Hilbert lattice: As discussed above, the events associated to quantum systems can be put in one to one correspondence with an orthomodular lattice: the one formed by the set of closed subspaces of a Hilbert space $\mathcal{H}$. They can be endowed with a lattice structure as follows [21]. The operation "$\vee$" is taken as the closure of 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 we denote by $\mathcal{P}(\mathcal{H})$ the set of closed subspaces. The order "$\leq$" is defined by subspace inclusion: we say that $\mathbb{S} \leq \mathbb{T}$, whenever $\mathbb{S} \subseteq \mathbb{T}$. The subspaces **0** and **1** play the role of the bottom and top elements of the lattice, since, for any subspace $\mathbb{S}$, we have $\mathbf{0} \leq \mathbb{S} \leq \mathbf{1}$.

Then, the algebraic structure $(\mathcal{P}(\mathcal{H}), \cap, \oplus, \neg, 0, 1)$ will be a complete bounded orthomodular lattice (which we denote simply by $\mathcal{P}(\mathcal{H})$). It is complete because the intersections and (the closure of) sums of arbitrary families of closed subspaces yields closed subspaces. It is bounded due to the existence of the top and bottom elements. It is orthomodular, because for any pair of subspaces $\mathbb{S}$, and $\mathbb{T}$, whenever we have $\mathbb{S} \leq \mathbb{T}$, then $\mathbb{S} \vee ((\mathbb{S})^{\perp} \wedge \mathbb{T}) = \mathbb{T}$ (see Appendix A for more details).

As closed subspaces are in one to one correspondence with projection operators, we 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 (A6) does not hold (see Appendix A). $\mathcal{P}(\mathcal{H})$ is modular if $\mathcal{H}$ is finite dimensional. If $\mathcal{H}$ is infinite dimensional, then $\mathcal{P}(\mathcal{H})$ is always orthomodular. Gleason's theorem (mentioned in the previous section) grants that, for $dim(\mathcal{H}) \geq 3$, quantum states can be considered as measures over the lattice of closed subspaces. This is a remarkable fact, since it implies that quantum probabilities are described by a very specific mathematical framework.

Any measurement context can be represented by an orthogonal basis of $\mathcal{H}$. It is easy to check that, by applying the lattice operations defined above to a fixed basis, we obtain a Boolean algebra. The cases of Hilbert spaces of dimension 2 and 3 are easy to check; however, this is true in general. It turns out that the whole lattice of subspaces can be described as a family of intertwined Boolean algebras (more about this below). For more discussion regarding the notion of "intertwined contexts", see [95].

The following table summarizes the main differences between Quantum and Kolmogorovian probabilities:

	Kolmogorov Probability	Quantum Probability
Lattice:	Σ	$\mathcal{P}(\mathcal{H})$
	(Boolean-algebra)	(orthomodular, non-Boolean)
States:	Measures over Σ	Measures over $\mathcal{P}(\mathcal{H})$
Events:	Subsets of Ω	Closed subspaces of $\mathcal{H}$

There is a geometrical underlying quantum probability: the one dimensional subspaces of a Hilbert space form a *projective geometry*. The higher dimensional subspaces are elements of the *projective lattice* associated with this geometry (see [38,39]).

Firefly Model: The firefly model [40] is used in quantum logic to show an example of a system that is not a full quantum model but has certain features that serve to illustrate what happens with quantum systems. It consists of a firefly that is freed inside a box. We are asked to perform an experiment to detect the location of the firefly but with constrains. We are only allowed to look at two different faces of the box (and we can only choose one on each run of the experiment). The first one is to measure face C_1 with three possible outcomes: the "firefly is detected on the left" (l), on the right (r), and "no-signal" (n) (which means that the light of the firefly was off).

The second possibility is to measure on face C_2, with the possibilities "firefly is on front" (f), "firefly is in the bottom" (b), and "no-signal". Notice that the "no-signal" outcome (n) is present in both experiments—this will be important soon. These constraints are, of course, silly, given that we can always look at every place in the box and detect the exact location of the firefly. However, they are thought off as an artificial measurement procedure that resembles what happens with quantum systems. If we choose context C_1, we can check whether the firefly is on the left, the right, or no signal. If we choose context C_2, we can check whether it is front, bottom, or no-signal. However, we cannot check both things in the same experiment, as happens with the position and momentum of a quantum system.

If we choose to measure in the face C_1, the three outcomes form an outcome set $\Omega_1 = \{(l), (r), (n)\}$. This gives rise to a Boolean algebra Σ_1, formed by all possible subsets of Ω_1. Each of these subsets represents an event, such as "the firefly is not detected on

the right" (which is represented by the set $\{(l), (n)\}$), and so on. A probabilistic state of the firefly —a throw in which we do not know the outcome a priori—will give a classical probability space $(\Omega_1, \Sigma_1, \mu_1)$. Similarly, we have a probability space $(\Omega_2, \Sigma_2, \mu_2)$ for the second option C_2, where $\Omega_2 = \{(t), (b), (n)\}$. Notice that $\Omega_1 \cap \Omega_2 = \{\mathbf{0}, (n), (n)', \mathbf{1}\}$. Since the event (n) belongs to both contexts of measurement, for the sake of consistency, we must impose $\mu_1((n)) = \mu_2(n)$.

Now, let $(n') = \{(l), (r)\}$, $(l') = \{(n), (r)\}$ and $(r') = \{(n), (l)\}$ (i.e., the set of theoretical complements of (n), (l) and (r), respectively). The Hasse diagram of Σ_1 is then given by
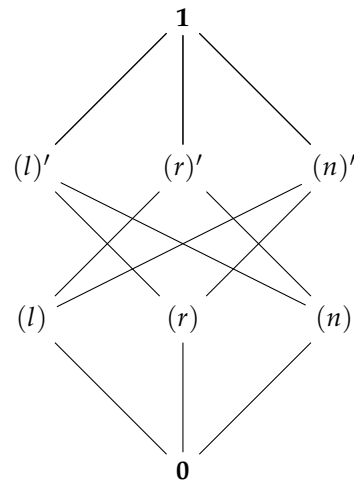

In the above diagram, a line joining two elements x and y, means that $x \leq y$ (i.e., the partial order is represented by the lines connecting the different elements). Thus, for example, $(l) \leq (r)'$ (which is equivalent to $\{(l)\} \subseteq \{(n), (l)\}$). The join of two elements is the least element that lies above both of them (with regards to the partial order). The conjunction is the greatest element that lies below both. Thus, for example, $(l) \vee (r) = (n)'$ (which means $\{(l)\} \cup \{(r)\} = (\{(n)\})^c$) and $(l) \wedge (r) = \mathbf{0}$ (which means $\{(l)\} \cap \{(r)\} = \varnothing$). A similar convention holds for the rest of the diagrams below.

Similarly, the Hasse diagram of Σ_2 is then given by
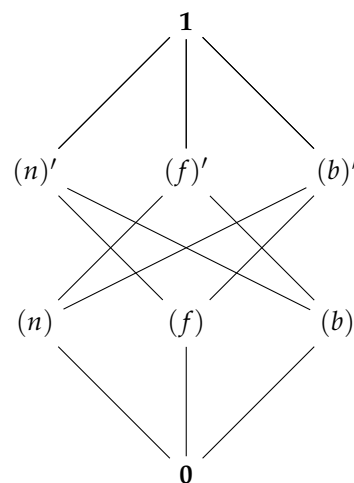
A direct check shows that Σ_1 and Σ_2 are Boolean algebras (also, Boolean lattices—see Appendix A). Now, we can join all possible events together, taking into account that $\Sigma_1 \cap \Sigma_2 = \{\mathbf{0}, (n), (n)', \mathbf{1}\}$. We obtain the following Hasse diagram:
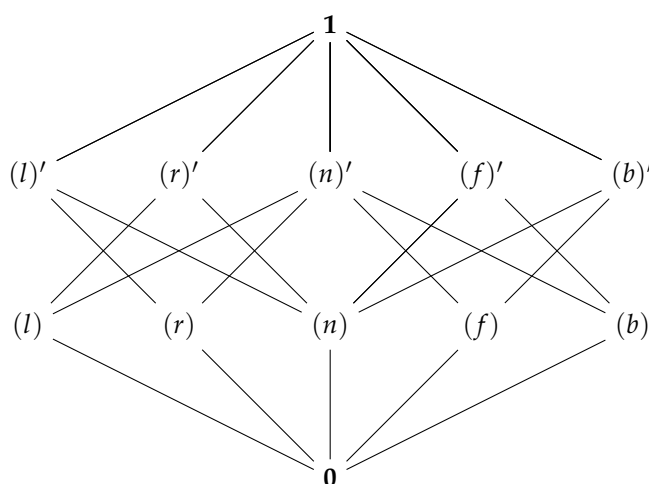
The above diagram defines a lattice $\mathcal{L}$, which—like Σ_1 and Σ_2—is non-distributive (and thus, non-Boolean). The lattice join of two given elements is the least element that lies above them, and the conjunction is the greatest element from below. The reader can check non-distributivity by inspection.

The Boolean algebras Σ_1 and Σ_2 are sublattices of $\mathcal{L}$. It is very important to remark that they contain an element in common: $\mathcal{L}$ can be seen a pasting of Σ_1 and Σ_2. In other words, $\mathcal{L}$ is formed from two Boolean subalgebras that are *intertwined*. The associated lattices of fully quantum models are just like that: they are formed by a collection of intertwined Boolean subalgebras—one for each context. The difference between the lattice of the firefly and the lattice of a three-dimensional quantum system is that there are infinitely many contexts for the latter, and thus the intertwining—for $dim(\mathcal{H}) \geq 3$—is much more complicated. This intricate algebraic structure associated to quantum systems lies at the core of the celebrated Kochen–Specker theorem [96] (which we discuss below).

The lattice of Q-bit: Given the incredible advances of quantum information theory in recent decades, the reader may wonder what the lattice of a q-bit looks like. It is the simplest quantum model conceivable. Suppose then that we are given a spin $\frac{1}{2}$ system. As is well known, the set of all possible states of a qubit is isomorphic to a sphere, namely, the *Bloch sphere* [91]. *Each pure state of a qubit corresponds to a one dimensional subspace of a two dimensional complex Hilbert space, and can be represented as a point in the surface of the Bloch sphere.*

A one dimensional subspace is called a *ray*. The different sets of objective properties, which are of the form "the particle has spin $\uparrow$ (or $\downarrow$) in direction $\hat{n}$", are represented by those rays (or, equivalently, by points on the surface of the sphere). Notice that each direction in space $\hat{n}$ defines two rays in the Hilbert space (represented by the projection operators $P_{\hat{n}}^{\uparrow} = | \uparrow \rangle \langle \uparrow |_{\hat{n}}$ and $P_{\hat{n}}^{\downarrow} = | \downarrow \rangle \langle \downarrow |_{\hat{n}}$).

The subspaces associated to $P_{\hat{n}}^{\uparrow}$ and $P_{\hat{n}}^{\downarrow}$ are orthogonal: this means, literally, that we must imagine them as orthogonal lines in the Hilbert space. As there are infinitely many directions in space, there are infinitely many such pairs of orthogonal events. All these events will be included in the lattice of a qubit.

In addition to all possible rays (associated to one dimensional subspaces), we also have two distinguished subspaces, represented by the events **0** (the null subspace of the Hilbert space), and **1** (the maximal subspace, which equals $\mathcal{H}$). Each one dimensional subspace contains **0** as a subspace, and is contained in **1**. If we chose a direction in space $\hat{n}$, consider the set $\mathcal{B}_{\hat{n}} = \{\mathbf{0}, P_{\hat{n}}^{\uparrow}, P_{\hat{n}}^{\downarrow}, \mathbf{1}\}$, and consider the above defined lattice operations for subspaces, we obtain a two-element Boolean algebra. All contexts of a qubit are of this form: each measurement direction $\hat{n}$ in space defines a two-elements Boolean algebra $\mathcal{B}_{\hat{n}}$. The Hasse diagram of a context represented by $\mathcal{B}_{\hat{n}}$ is then given by
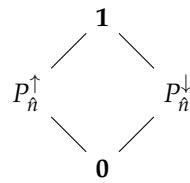
Thus, the Hasse diagram of a q-bit will have the form:
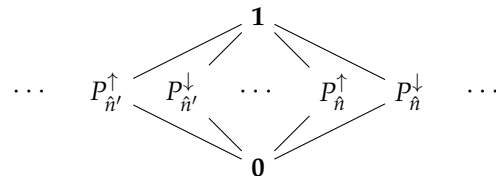
where $\hat{n}$, $\hat{n}'$, etc., define different directions in space. The dots represent the infinitely many other Boolan algebras associated with all possible directions in space. Again, we obtain a lattice, which is non-distributive. In this example, $\mathcal{B}_{\hat{n}} \cap \mathcal{B}_{\hat{n}'} = \{\mathbf{0}, \mathbf{1}\}$ whenever $\hat{n}$ and $\hat{n}'$ define different directions. Thus, only the top and bottom elements are shared by the Boolean subalgebras. Thus, this example is degenerated, since there is no (non-trivial) intertwining between the different Boolean algebras associated to the measurement contexts. In the following example, we will consider a higher dimensional example, for which the intertwining is highly non-trivial.

Kochen–Specker theorem (in a four dimensional model): A nice example of how the different contexts of a quantum system are intertwined was presented in [97] (of course, for the original version of the Kochen–Specker theorem, the reader is referred to [96]). Given a four-dimensional quantum system, each measurement context has four possible outcomes. Each one of them is mathematically represented by a one dimensional subspace of a four-dimensional Hilbert space.

Each one dimensional subspace is generated by a vector $\hat{v}$. Let us then represent the outcome given by the vector $\hat{v}$ by the projection operator $P_{\hat{v}}$ (which is the projection operator that projects into the subspace generated by $\hat{v}$). Then, each measurement context is represented by four projection operators, say $P_{\hat{v}_1}$ $P_{\hat{v}_2}$, $P_{\hat{v}_3}$ and $P_{\hat{v}_4}$. These are all orthogonal, because they represent mutually exclusive outcomes (you cannot have, in the same measurement context, two different outputs at the same time). As in the qubit case, by using the Hilbert lattice operations, these projections generate a Boolean algebra with four atoms (that has 2^4 elements).

Thus, each measurement context has an associated Boolean algebra with sixteen elements. When a quantum state is prepared, the probabilities assigned to each context can be thought off as Kolmogorovian. It also happens that $P_{\hat{v}_1} + P_{\hat{v}_2} + P_{\hat{v}_3} + P_{\hat{v}_4} = \mathbf{1}$ (which, using the Hilbert lattice operations, reads $\bigvee_i P_{\hat{v}_i} = \mathbf{1}$). Now, consider a family of nine different measurement contexts, giving place to the equations:

$$
\begin{aligned}
P_{0,0,0,1} + P_{0,0,1,0} + P_{1,1,0,0} + P_{1,-1,0,0} &= \hat{1} \\
P_{0,0,0,1} + P_{0,1,0,0} + P_{1,0,1,0} + P_{1,0,-1,0} &= \hat{1} \\
P_{1,-1,1,-1} + P_{1,-1,-1,1} + P_{1,1,0,0} + P_{0,0,1,1} &= \hat{1} \\
P_{1,-1,1,-1} + P_{1,1,1,1} + P_{1,0,-1,0} + P_{0,1,0,-1} &= \hat{1} \\
P_{0,0,1,0} + P_{0,1,0,0} + P_{1,0,0,1} + P_{1,0,0,-1} &= \hat{1} \\
P_{1,-1,-1,1} + P_{1,1,1,1} + P_{1,0,0,-1} + P_{0,1,-1,0} &= \hat{1} \\
P_{1,1,-1,1} + P_{1,1,1,-1} + P_{1,-1,0,0} + P_{0,0,1,1} &= \hat{1} \\
P_{1,1,-1,1} + P_{-1,1,1,1} + P_{1,0,1,0} + P_{0,1,0,-1} &= \hat{1} \\
P_{1,1,1,-1} + P_{-1,1,1,1} + P_{1,0,0,1} + P_{0,1,-1,0} &= \hat{1},
\end{aligned}
\tag{20}
$$

Each line of Equation (20) below represents a different measurement context. The subindexes represent non-normalized vectors in the four-dimensional Hilbert space (each vector defines a ray, and represents a measurement outcome in that context). Each measurement context has four orthogonal projections, that represent the outcomes of a projective measurement (and, as such, add up to the identity operator). The contexts are chosen in such a way that each pair of them shares one element in common.

This represents the intertwining of the Boolean algebras associated to the contexts since, for example, the event represented by $P_{0,0,0,1}$ belongs to the first and second contexts. Similarly, the event $P_{0,0,1,0}$ belongs to the first and the fifth contexts. The family of contexts of this example is chosen in such a way that each event belongs to exactly two different contexts. Thus, since there are nine equations, there are eighteen different events in total.

The Boolean algebras associated to the nine contexts are related in a non-trivial way (since the intersection of any two of them is strictly greater than $\{\mathbf{0}, \mathbf{1}\}$). In order to illustrate the Kochen–Specker theorem, let try to assign truth values to each of these events, which can be represented as 0 vs. 1 assignments to the different outcomes of the experiments. Thus, for example, we can assign 1 (true) to $P_{0,0,0,1}$, or 0 false, and proceed similarly to the other events.

We represent this by a function ν: $\nu(P_{0,0,0,1}) = 1$, $\nu(P_{0,0,1,0}) = 0$, etc. A truth value assignment would mean that each possible experiment outcome has a definite value previous to measurement. This is related to asking about the existence of a *dispersion free state*, that is, a state that only assigns the probabilities zero and one to all possible outcomes. Thus, the function ν must satisfy one condition: given that all the outcomes in a context are mutually exclusive, the function must be defined in such a way that there are no two truth value assignments in the same context.

Thus, for example, if we assign the truth value 1 to $P_{0,0,0,1}$ (i.e., $\nu(P_{0,0,0,1}) = 1$), then, all other members of that context must have the truth value 0 assigned ($\nu(P_{0,0,1,0}) = \nu(P_{1,1,0,0}) = \nu(P_{1,-1,0,0}) = 0$). Equation (20) implies that the valuations must satisfy $\sum_i \nu(P_i) = 1$ on each line (this is known as the FUNC condition in the literature; see the discussion and references in [98]). We must assume that the valuations preserve their values from context to context (if we assign a certain truth value to a projection in a given context, we must use that same value when it appears in a different context).

However, it is easy to check that such a compatible truth value assignment is not possible. The reason is as follows. There are nine equations and eighteen events. If we sum all equations (of the form $\sum_i \nu(P_i) = 1$), on the right, we obtain an odd number (nine), and on the left, we obtain an even number, since there is an even number of ones. However, this is impossible. The non-existence of such a truth value assignment shows one of the most important implications of the intertwining between the Boolean algebras associated to the contexts.

This is known as the Kochen–Specker theorem (see [97] for details). This example illustrates clearly how the Boolean algebras of events associated to quantum systems are intertwined and how this complex structure gives place to interpretational issues. As is well known, the Kochen–Specker theorem is a cornerstone in the discussions about the foundations of quantum mechanics (see, for example [98] and the references therein).

3.4. 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 the *quantal effect*. This notion can be generalized to arbitrary statistical theories. The generalization of the notion of PVM (which is based on projections) to an observable based on effects is called a *positive operator valued measure* (POVM) [99–105] and, in QM, will be represented by a mapping ($\mathcal{B}(\mathcal{H})$ stands for the set of bounded operators in $\mathcal{H}$).

$$E : B(\mathbb{R}) \to \mathcal{B}(\mathcal{H}). \tag{21a}$$

such that

$$E(\mathbb{R}) = \mathbf{1} \tag{21b}$$

$$E(B) \geq 0, \text{ for any } B \in B(\mathbb{R}) \tag{21c}$$

$$E(\cup_j(B_j)) = \sum_j E(B_j), \text{ for any disjoint familly } B_j \tag{21d}$$

The reader should compare Equations (21a)–(21d) with (5a)–(5e) and (16a)–(16e). 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 orthogonal projections. It is the most general formulation of the description of a measurement in the framework of quantum physics. Positive operators E satisfying $0 \leq E \leq \mathbf{1}$ are called *effects* and generate an *effect algebra* [99,102]). We denote this algebra by $E(\mathcal{H})$. It is also important to remark that POVMs can be associated to fuzzy measurements (and thus with fuzzy sets; see [104,106]).

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. Thus, given a Borel set B, we have:

$$E(B) : \mathcal{C} \to [0, 1] \tag{22a}$$

$$\rho \mapsto \mathrm{tr}(E(B)\rho) \tag{22b}$$

for every Borel set B. This will be relevant in certain generalizations of quantum probabilities, which we will discuss below.

4. Generalization to Orthomodular Lattices

In the algebraic formulation of relativistic quantum theory, there appear algebras that are different from the ones used in non-relativistic QM [8]. In the non-relativistic case, the algebra $\mathcal{B}(\mathcal{H})$ of all bounded operators acting on a separable Hilbert space generates— via the spectral theorem—all possible observables. However, the study of quantum systems with infinitely many degrees of freedom revealed that other algebras are needed. Murray and von Neumann provided a classification of these algebras, which are called Type I, Type II, and Type III. For the non-relativistic case with finitely many degrees of freedom, it suffices to use Type I factors. However, in the general case, Type II and Type III factors appear.

The existence of different algebraic models of quantum theories suggests that, in principle, one could conceive more general probabilistic models than those of standard QM. We describe here a possible generalization, based in orthomodular lattices. Let $\mathcal{L}$ be an arbitrary orthomodular lattice (standing for the lattice of all possible empirical events of a given model). Then, we define

$$s : \mathcal{L} \to [0; 1], \tag{23a}$$

such that:

$$s(\mathbf{1}) = 1. \tag{23b}$$

and, for a denumerable and pairwise orthogonal family of events E_j,

$$s(\sum_j E_j) = \sum_j s(E_j). \tag{23c}$$

If we put $\mathcal{L} = \Sigma$ and $\mathcal{L} = \mathcal{P}(\mathcal{H})$ in Equations (23a)–(23c), we recover the Kolmogorovian and quantum cases, respectively. For a discussion on the conditions under which measures those defined in Equations (23a)–(23c) are well defined, see [37], Chapter 11. The fact that projection operators of arbitrary von Neumann algebras define orthomodular lattices [12], shows that the above generalization includes many examples of interest (in addition to classical statistical mechanics and standard QM).

Notice again that the set of all possible measures satisfying (23a)–(23c) is convex. This opens the door to a further generalization of probabilistic models based on *convex sets*, that we discuss in the next Section.

The states defined in Equations (23a)–(23c) define Kolmogorovian probabilities when restricted to maximal Boolean subalgebras of $\mathcal{L}$. Denote by **B** to the set of all possible Boolean subalgebras of $\mathcal{L}$. It is possible to consider **B** as a pasting of its maximal Boolean subalgebras (see, for example [107] and the discussions posed in [20,108]):

$$\mathcal{L} = \bigvee_{\mathcal{B} \in \mathbf{B}} \mathcal{B}. \tag{24}$$

The decomposition represented by Equation (24) implies that a state defined as a measure over an orthomodular lattice can be considered as a pasting of Kolmogorovian probabilities. If there is only one maximal Boolean subalgebra, then the whole $\mathcal{L}$ has to be Boolean, and thus we recover a Kolmogorovian model. In theories that display contextuality, such as standard QM [20,94,108], there will be more than one empirical context, and thus the above decomposition will not be trivial.

The representation of observables in this setting can be made as follows (we follow [43] here).

Definition 1. *A c-morphism is a one to one map $\alpha : \mathcal{L}_1 \longrightarrow \mathcal{L}_2$ between orthocomplemented complete lattices $\mathcal{L}_1$ and $\mathcal{L}_2$ such that*

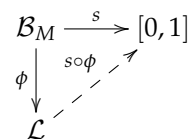$$\alpha(\bigvee_i a_i) = \bigvee_i \alpha(a_i) \tag{25a}$$

$$a \perp b \Longrightarrow \alpha(a) \perp \alpha(b) \tag{25b}$$

$$\alpha(\mathbf{1}_1) = \alpha(\mathbf{1}_2) \tag{25c}$$

Given a physical system whose event lattice is given by $\mathcal{L}$, an *observable* can be defined as a c-morphism from a Boolean lattice $\mathcal{B}$ into $\mathcal{L}$:

Definition 2 (Observable). *An observable of a physical system whose event lattice is $\mathcal{L}$ and that takes its values in the outcome set M will be a c-morphism ϕ from a Boolean algebra $\mathcal{B}_M$ of subsets of M, to a Boolean subalgebra $\Sigma_\phi \subseteq \mathcal{L}$.*

The following diagram illustrates what happens when we compose a generalized state (given by Equations (23a)–(23c)) and an observable (as in Definition 2):

$$
\begin{array}{ccc}
\mathcal{B}_M & \xrightarrow{\ s\ } & [0,1] \\
{\scriptstyle \phi} \downarrow & {\scriptstyle s \circ \phi} \nearrow & \\
\mathcal{L} & &
\end{array}
$$

The composition $s \circ \phi : \mathcal{B}_M \longrightarrow [0,1]$ defines a Kolmogorovian probability (i.e., a measure satisfying Equations (1a)–(1c)).

Let us now compare compare Equations (5a)–(5e) and (16a)–(16e) with Definitions 1 and 2. By looking at the definition of PVM (Equations (16a)–(16e)), it is easy to recognize that a PVM is a c-morphism between the set of Borel subsets of $\mathcal{B}(\mathbb{R})$ and the Boolean algebra generated by its image projections. According to the above definition of observable, one can quickly realize that any Boolean subalgebra of $\mathcal{L}$ will determine an observable (more properly, a family of observables up to rescaling). For the classical case, by looking again at the "important remark" of Section 2.2 (Equations (5a)–(5e)), we realize that a classical random variable also satisfies the general definition of observable given in Definition 2.

5. Convex Operational Models

In the previous section, we demonstrated that the set of states defined over an arbitrary orthomodular lattice is convex. This approach contains the quantum and classical state spaces as particular cases. Thus, it seems very natural to attempt to define generalized probabilistic models by appealing to convex sets.

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 *Convex Operational Models* (COM) approach. In this section, we concentrate on elementary notions of COM's, and we refer the reader to [57] for an excellent presentation of the subject. The approach based on convex sets results as more general than the one based in orthomodular lattices (i.e., the latter can be included as particular cases of the COM approach).

If the state space of a given probabilistic theory is given by the set **S**, 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

$$p : X \times \mathbf{S} \mapsto [0, 1]$$
$$(x, s) \to p(x, s) \tag{26}$$

To each outcome $x \in X$ and state $s \in \mathbf{S}$, 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 $(\mathbf{S}, p(\cdot, \cdot), X)$ is assigned for each system of any probabilistic theory [17]. Thinking of s as a variable, we obtain a mapping $s \mapsto p(\cdot, s)$ from $\mathbf{S} \to [0, 1]^X$. This implies that all the states of **S** can be identified with maps, which generates a canonical vector space. Their closed convex hull forms a new set $\mathcal{S}$ representing all possible probabilistic mixtures (convex combinations) of states in **S**. Given an arbitrary $\alpha \in \mathcal{S}$ and any outcome $x \in X$, we can define an affine evaluation-functional $f_x : \mathcal{S} \to [0, 1]$ in a canonical way by $f_x(\alpha) = \alpha(x)$.

More generally, we can consider any affine functional $f : \mathcal{S} \to [0, 1]$ as representing a measurement outcome, and thus use $f(\alpha)$ to represent the probability for that outcome in state α. We will call $A(\mathcal{S})$ to the space of all affine functionals. Due to the fact that QM is also a probabilistic theory, it follows that it can be included in the general framework described above (we denoted its convex set of states by $\mathcal{C}$ in Section 3.2). In QM, affine functionals defined as above are called *effects* (and are coincident with the constituents of POVM's as defined in Section 3.4). The generalized probabilistic models defined in Section 4 fall naturally into the scope of the COM approach, given that their state spaces are convex sets.

We saw that a probability $a(\omega) \in [0, 1]$ is well defined for any state $\omega \in \mathcal{S}$ 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 $\omega \in \mathcal{S}$. 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 space of affine functionals). A (discrete) measurement will be represented by a set of effects $\{a_i\}$ such that $\sum_i a_i = u$. $\mathcal{S}$ can be naturally embedded in the dual space $A(\mathcal{S})^*$ using the map

$$\omega \mapsto \hat{\omega}$$
$$\hat{\omega}(a) := a(\omega) \tag{27}$$

Let $V(\mathcal{S})$ be the linear span of $\mathcal{S}$ in $A(\mathcal{S})^*$. Then, it is reasonable to consider $\mathcal{S}$ finite dimensional if and only if $V(\Omega)$ is finite dimensional. For the sake of simplicity, we restrict ourselves to this case (and to compact spaces). As is well known, this implies that $\mathcal{S}$ 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 $\mathcal{C}$, and correspond to one dimensional projections in the Hilbert space).

It can be shown that, for finite dimensions d, a system will be classical if and only if it is a simplex (a simplex is the convex hull of $d + 1$ linearly independent pure states). 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 state 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 that are stable under mixing and purification. This is means that a convex subset F is a face if, each time that

$$x = \lambda x_1 + (1 - \lambda)x_2, \quad 0 \leq \lambda \leq 1, \tag{28}$$

and then $x \in F$ if and only if $x_1 \in F$ and $x_2 \in F$ [91]. The set of faces of a convex set forms 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 $\mathcal{P}(\mathcal{H})$ [37,91]. For a general model, the lattice of faces may fail to be suitably orthocomplemented [37] (and thus the COM approach is more general than the one based in orthomodular lattices).

Let us turn now to compound systems. Given a compound system, its components will have state spaces $\mathcal{S}_A$ and $\mathcal{S}_B$. Let us denote the joint state space by $\mathcal{S}_{AB}$. It is reasonable to identify $\mathcal{S}_{AB}$ with the linear span of $(V(\mathcal{S}_A) \otimes V(\mathcal{S}_B))$ [57]. Then, a maximal tensor product state space $\mathcal{S}_A \otimes_{max} \mathcal{S}_B$ can be defined as one that contains all bilinear functionals $\varphi : A(\mathcal{S}_A) \times A(\mathcal{S}_B) \longrightarrow \mathbb{R}$ 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 largest set of states in $(A(\mathcal{S}_A) \otimes A(\mathcal{S}_B))^*$, which assigns probabilities to all product-measurements. The minimal tensor product state space $\mathcal{S}_A \otimes_{min} \mathcal{S}_B$ 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

$$\omega_A \otimes \omega_B(a, b) = \omega_A(a)\omega_B(b), \tag{29}$$

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

$$\mathcal{S}_A \otimes_{min} \mathcal{S}_B \subseteq \mathcal{S}_A \otimes \mathcal{S}_B \subseteq \mathcal{S}_A \otimes_{max} \mathcal{S}_B \tag{30}$$

As expected, for classical compound systems (because of the absence of entangled states), we have $\mathcal{S}_A \otimes_{min} \mathcal{S}_B = \mathcal{S}_A \otimes_{max} \mathcal{S}_B$. In the quantum case, we have strict inclusions in (30): $\mathcal{S}_A \otimes_{min} \mathcal{S}_B \subsetneq \mathcal{S}_A \otimes \mathcal{S}_B \subsetneq \mathcal{S}_A \otimes_{max} \mathcal{S}_B$. The general definition of a separable state in an arbitrary COM is made in analogy with that of [109], i.e., as one that can be written as a convex combination of product states [58,62] (see also [64] for a generalization):

Definition 3. *A state $\omega \in \mathcal{S}_A \otimes \mathcal{S}_B$ will be called* separable *if there exist $p_i \in \mathbb{R}_{\geq 0}$, $\omega_A^i \in \mathcal{S}_A$ and $\omega_B^i \in \mathcal{S}_B$ such that*

$$\omega = \sum_i p_i \omega_A^i \otimes \omega_B^i, \quad \sum_i p_i = 1 \tag{31}$$

If $\omega \in \mathcal{S}_A \otimes \mathcal{S}_B$ is not separable, then it will be reasonably called *entangled* [91,110]. As expected, entangled states exist only if $\mathcal{S}_A \otimes \mathcal{S}_B$ is strictly greater than $\mathcal{S}_A \otimes_{min} \mathcal{S}_B$.

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 $\mathcal{S}$. 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.

It is important to remark that many informational techniques, such as the MaxEnt method, can be suitably generalized to arbitrary probabilistic models [111,112]. In a similar vein, quantum information theory could be considered as a particular case of a generalized information theory [108].

6. Cox's Method Applied To Physics

Now, we review a relatively recent approach to the probabilities appearing in QM that uses distributive lattices. A novel derivation of Feynman's rules for quantum mechanics was presented in [29,34]. There, 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 3.2 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 $M_1, \ldots, M_n$ on a quantum system, yielding results $m_1, m_2, \ldots, m_n$, 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

$$P(A) = Pr(m_n, \ldots, m_2 | m_1) \tag{32}$$

representing the probability of obtaining outcomes $m_2, \ldots, m_n$ conditional upon obtaining m_1.

Let us see how this works with a concrete example in which the m_i's have two possible values, 0 and 1. Then, $A_1 = [0, 1, 1]$ and $A_2 = [0, 0, 1, 1]$ represent the measurement sequences of three and four measurements, respectively. Here, $P(A_1) = Pr(1, 1 | 0)$ 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 $\tilde{M}_2$ is created. Thus, a possible sequence obtained by the replacement of M_2 by $\tilde{M}_2$ could be $[1, (0, 1), 1]$. Analogous constructions can be done for other measurements. In this way, an operation can be defined for the sequences:

$$[m_1, \ldots, (m_i, m_i'), \ldots, m_n] := [m_1, \ldots, m_i, \ldots, m_n] \vee [m_1, \ldots, m_i', \ldots, m_n] \tag{33}$$

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

$$[m_1, \ldots, m_j, \ldots, m_n] := [m_1, \ldots, m_j] \cdot [m_j, \ldots, m_n] \tag{34}$$

Notice that, in the above equation, the last measurement and outcome of the first sequence must be the same as the first measurement and outcome of the second. With these operations at hand, it is easy to show that, if A, B, and C are measuring sequences, then

$$A \vee B = B \vee A \tag{35a}$$

$$(A \vee B) \vee C = A \vee (B \vee C) \tag{35b}$$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \tag{35c}$$

$$(A \vee B) \cdot C = (A \cdot C) \vee (B \cdot C) \tag{35d}$$

$$C \cdot (A \vee B) = (C \cdot A) \vee (C \cdot B), \tag{35e}$$

Equations (35a)–(35e) show explicitly that "∨" is commutative and associative, "·" is associative, and that there is right- and left-distributivity of "·" over "∨".

Equations (35a)–(35e) 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 [34]. However, at this step, a crucial assumption is made: each measuring sequence will be represented by a pair of real numbers [34]. This assumption is justified in [34] by appealing to Bohr's complementarity principle.

If measuring sequences A, B, etc. induces pairs of real numbers $\mathbf{a}$, $\mathbf{b}$, etc., then, due to Equations (35a)–(35e), the associated pairs should satisfy

$$\mathbf{a} \vee \mathbf{b} = \mathbf{b} \vee \mathbf{a} \tag{36a}$$

$$(\mathbf{a} \vee \mathbf{b}) \vee \mathbf{c} = \mathbf{a} \vee (\mathbf{b} \vee \mathbf{c}) \tag{36b}$$

$$(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c}) \tag{36c}$$

$$(\mathbf{a} \vee \mathbf{b}) \cdot \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) \vee (\mathbf{b} \cdot \mathbf{c}) \tag{36d}$$

$$\mathbf{c} \cdot (\mathbf{a} \vee \mathbf{b}) = (\mathbf{c} \cdot \mathbf{a}) \vee (\mathbf{c} \cdot \mathbf{b}) \tag{36e}$$

The reader can easily verify that Equations (36a)–(36e) are satisfied by the field of complex numbers (provided that the operations are interpreted as sum and product of complex numbers). How can we be assured that complex numbers are the only field that satisfies Equations (36a)–(36e)? In order to single out complex numbers among other possible fields, additional assumptions must be added, namely, pair symmetry, additivity, and symmetric bias (see [29,34] 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 (35a)–(35e).

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 [24,25] and the one of A. N. Kolmogorov [3]. The Kolmogorovian approach can be generalized in order to include non-Boolean models, as we have shown in Section 4. 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 [19,27].

Generalized Probability Calculus Using Cox's Method

As we have seen in Section 2, Cox studied the functions defined over a distributive lattices and derived classical probabilities. In [27], it is shown that if the lattice is assumed to be non-distributive, the properties of QP described in Section 3.2 can be derived by applying a variant of Cox's method as follows (see [27] for details). Suppose that the propositions of our system are represented by the lattice of elementary tests of QM, i.e., 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 Equations (17a)–(17c). 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 "∨" 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

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

with F a function to be determined. If a third proposition R is added, following a similar procedure to that of Cox, we obtain for "$P \vee P \vee R$" the following functional equation

$$F(F(s(P), s(Q)), s(R)) = F(s(P), F(s(Q), s(R))). \tag{38}$$

The above equation can be solved up to rescaling [30,31,33,113], and we find

$$s(P \vee Q) = s(P) + s(Q). \tag{39}$$

whenever $P \perp Q$. It can be shown that, for any finite family of orthogonal projections P_j, $1 \leq j \leq n$ [27]:

$$s(\bigvee_{j=1}^{\infty} P_j) = \sum_{j=1}^{\infty} s(P_j), \tag{40}$$

and we recover condition (23c) of the axioms of quantum probability. By exploiting the properties of the orthogonal complement acting on subspaces, it can also be shown [27] that

$$s(P^{\perp}) = 1 - s(P), \tag{41}$$

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 (23b). In this way, it follows that Cox's method applied to the non-distributive lattice $\mathcal{P}(\mathcal{H})$ yields the same probability theory as the one provided by Equations (17a)–(17c) for the quantum case.

What happens if Cox's method is applied to an arbitrary atomic orthomodular complete 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 [27], it is shown that, under these rather general assumptions, in *any* atomic orthomodular lattice and for any orthogonal denumerable family $\{a_i\}_{i \in \mathbb{N}}$, s must satisfy (up to rescaling)

$$s(\bigvee \{a_i\}_{i \in \mathbb{N}}) = \sum_{i=1}^{\infty} s(a_i) \tag{42a}$$

$$s(\neg a) = 1 - s(a) \tag{42b}$$

$$s(\mathbf{0}) = 0. \tag{42c}$$

In this way, a generalized probability theory is derived (as in (17a)–(17c)). Equations (42a)–(42c) 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

$$(a \wedge b) \vee (a \wedge \neg b) < a, \tag{43}$$

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

In the non-Boolean setting of QM, von Neumann's entropy (VNE) plays a similar role to that of Shannon's in Cox approach [20]. This allows us to interpret the VNE as a natural measures of information for an experimenter who deals with a contextual event structure.

8. Conclusions

We presented a new approach for probabilities appearing in QM. While there exist (at least) 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 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.

These formal frameworks do not exhaust the philosophical debate around the existence or not of a well-defined notion of QP; notwithstanding, the extension of Cox's method to the non-distributive case, as well as 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. According to this interpretation, a rational agent is confronted with a particular event structure. To fix ideas, suppose that the agent

is confronted with a physical system, and that the agent has to perform experiments and determine degrees of belief about their possible outcomes.

- If the lattice of events that the agent is facing is Boolean (as in Cox's approach), then, the measures of degree of belief will obey laws equivalent to those of Kolmogorov.
- On the contrary, if the state of affairs that the agent must face presents contextuality (as in standard quantum mechanics), the measures involved must be non-Kolmogorovian [27].
- Random variables and information measures [20] will be the natural generalizations of the classical case if the event structure is not classical. A similar observation holds for the application of the MaxEnt method [111,112].

Our approach allows for a natural justification of the peculiarities arising in quantum phenomena from the standpoint of a Bayesian approach. In particular, quantum information theory could be considered as a non-Kolmogorovian extension of Shannon's theory [108]. Our approach can be considered as an alternative step to address Hilbert's problem for the case of probability theory in QM: the development of an axiomatization endowed with a clear and natural interpretation of the notions involved.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: .The authors declare no conflict of interest.

Appendix A. 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 all $x, y, z \in \mathcal{L}$, the following equations are satisfied.

$$x \vee x = x \, , \; x \wedge x = x \quad \text{(idempotence)} \tag{A1a}$$

$$x \vee y = y \vee x \, , \; x \wedge y = y \wedge x \quad \text{(commutativity)} \tag{A1b}$$

$$x \vee (y \vee z) = (x \vee y) \vee z \, , \; x \wedge (y \wedge z) = (x \wedge y) \wedge z \quad \text{(associativity)} \tag{A1c}$$

$$x \vee (x \wedge y) = x \wedge (x \vee y) = x \quad \text{(absorption)} \tag{A1d}$$

If the extra relationships

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \quad \text{(distributivity 1)} \tag{A2a}$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \quad \text{(distributivity 2)} \tag{A2b}$$

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 "$x \wedge y$" is 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 **1** and a least element **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 "$\neg$" satisfying:

$$\neg(\neg(a)) = a \tag{A3a}$$

$$a \leq b \longrightarrow \neg b \leq \neg a \tag{A3b}$$

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

$$a \vee \neg a = \mathbf{1} \tag{A3c}$$

$$a \wedge \neg a = \mathbf{0} \tag{A3d}$$

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 ($a \perp b$) if $a \leq \neg b$. Following [71], we define an *orthomodular lattice* as an ortholattice satisfying the orthomodular law:

$$a \leq b \implies a \vee (\neg a \wedge b) = b \tag{A4}$$

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

$$a \leq b \implies a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \tag{A5}$$

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

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \tag{A6}$$

Thus, a *Boolean lattice* is a **complemented distributive lattice**. We use the terms *Boolean lattice* and *Boolean algebra* interchangeably.

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 $0 < y < x$. $\mathcal{L}$ is 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.

Boolean	**Modular**	**Orthomodular**
$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$	$a \leq b \implies a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$	$a \leq b \implies a \vee (\neg a \wedge b) = b$

References

1. Hilbert, D. Mathematical problems. *Bull. Am. Math. Soc.* **1902**, *8*, 437–479. [CrossRef]
2. Rocchi, P. *Janus-Faced Probability*; Springer: New York, NY, USA, 2014.
3. Kolmogorov, A. *Foundations of the Theory of Probability: Second English Edition*; Courier Dover Publications: Mineola, NY, USA, 2018.
4. Lacki, J. The early axiomatizations of quantum mechanics: Jordan, von Neumann and the continuation of Hilbert's program. *Arch. Hist. Exact Sci.* **2000**, *54*, 279–318. [CrossRef]
5. Von Neumann, J. *Mathematical Foundations of Quantum Mechanics*; Princeton University Press: Princeton, NJ, USA, 2018.
6. Haag, R. *Local Quantum Physics: Fields, Particles, Algebras*; Springer Science & Business Media: New York, NY, USA, 2012.
7. Streater, R.F.; Wightman, A.S. *PCT, Spin and Statistics, and All That*; Princeton University Press: Princeton, NJ, USA, 2016.
8. Halvorson, H.; Müger, M. *Handbook of the Philosophy of Physics, Chapter Algebraic Quantum Field Theory*; North Holland: Amsterdam, The Netherlands, 2006.
9. Bratteli, O.; Robinson, D.W. *Operator Algebras and Quantum Statistical Mechanics: Volume 1: C*-and W*-Algebras. Symmetry Groups. Decomposition of States*; Springer Science & Business Media: New York, NY, USA, 2012.

10. Feynman, R.P. The concept of probability in quantum mechanics. In Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 31 July–12 August 1950; University of California Press: Berkeley, CA, USA, 1951; pp. 533–541
11. Fuchs, C.A.; Schack, R. Quantum-bayesian coherence. *Rev. Mod. Phys.* **2013**, *85*, 1693. [CrossRef]
12. Rédei, M.; Summers, S.J. Quantum probability theory. *Stud. Hist. Philos. Sci. Part B Stud. Hist. Philos. Mod. Phys.* **2007**, *38*, 390–417. [CrossRef]
13. Davies, E.B.; Lewis, J.T. An operational approach to quantum probability. *Commun. Math. Phys.* **1970**, *17*, 239–260. [CrossRef]
14. Srinivas, M. Foundations of a quantum probability theory. *J. Math. Phys.* **1975**, *16*, 1672–1685. [CrossRef]
15. Mackey, G.W. Quantum mechanics and Hilbert space. *Am. Math. Mon.* **1957**, *64*, 45–57. [CrossRef]
16. Gudder, S.P. *Stochastic Methods in Quantum Mechanics*; Courier Corporation: Mineola, NY, USA, 2014.
17. Mackey, G.W. *Mathematical Foundations of Quantum Mechanics*; Courier Corporation: Mineola, NY, USA, 2013.
18. Kujala, J.V.; Dzhafarov, E.N.; Larsson, J.Å. Necessary and sufficient conditions for an extended noncontextuality in a broad class of quantum mechanical systems. *Phys. Rev. Lett.* **2015**, *115*, 150401. [CrossRef]
19. Holik, F. Logic, geometry and probability theory. *SOP Trans. Theor. Phys.* **2014**, *1*, 128–137. [CrossRef]
20. Holik, F.; Plastino, A.; Sáenz, M. Natural information measures in Cox'approach for contextual probabilistic theories. *arXiv* **2015**, arXiv:1504.01635.
21. Rédei, M. *Quantum Logic in Algebraic Approach*; Springer Science & Business Media: New York, NY, USA, 2013; Volume 91.
22. Hamhalter, J. *Quantum Measure Theory*; Springer Science & Business Media: New York, NY, USA, 2013; Volume 134.
23. De Finetti, B. *Teoria delle Probabilità. Volume Primo*; Einaudi: Torino, Italy, 1970.
24. Cox, R.T. The algebra of probable inference. *Am. J. Phys.* **1963**, *31*, 66–67. [CrossRef]
25. Cox, R.T. Probability, frequency and reasonable expectation. *Am. J. Phys.* **1946**, *14*, 1–13. [CrossRef]
26. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: Cambridge, UK, 2003.
27. Holik, F.; Sáenz, M.; Plastino, A. A discussion on the origin of quantum probabilities. *Ann. Phys.* **2014**, *340*, 293–310. [CrossRef]
28. Caticha, A. Consistency, amplitudes, and probabilities in quantum theory. *Phys. Rev. A* **1998**, *57*, 1572. [CrossRef]
29. Goyal, P.; Knuth, K.H.; Skilling, J. Origin of complex quantum amplitudes and Feynman's rules. *Phys. Rev. A* **2010**, *81*, 022109. [CrossRef]
30. Knuth, K.H. Deriving laws from ordering relations. *AIP Conf. Proc.* **2004**, *707*, 204–235.
31. Knuth, K.H. Measuring on lattices. *AIP Conf. Proc.* **2009**, *1193*, 132–144.
32. Knuth, K.H. Valuations on lattices and their application to information theory. In Proceedings of the 2006 IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, 16–21 July 2006; pp. 217–224.
33. Knuth, K.H. Lattice duality: The origin of probability and entropy. *Neurocomputing* **2005**, *67*, 245–274. [CrossRef]
34. Goyal, P.; Knuth, K.H. Quantum theory and probability theory: Their relationship and origin in symmetry. *Symmetry* **2011**, *3*, 171–206. [CrossRef]
35. Birkhoff, G.; Von Neumann, J. The logic of quantum mechanics. *Ann. Math.* **1936**, *37*, 823–843. [CrossRef]
36. Dalla Chiara, M.L.; Giuntini, R.; Greechie, R. *Reasoning in Quantum Theory: Sharp and Unsharp Quantum Logics*; Springer Science & Business Media: New York, NY, USA, 2013; Volume 22.
37. Beltrametti, E.G.; Cassinelli, G. *The Logic of Quantum Mechanics: Volume 15*; Cambridge University Press: Cambridge, UK, 2010; Number 15.
38. Varadarajan, V.S. *Geometry of Quantum Theory*; Springer: New York, NY, USA, 1968; Volume 1.
39. Varadarajan, V. *Geometry of Quantum Theory, Volume II*; D. Van Nostrand Company: New York, NY, USA, 1970.
40. Svozil, K. *Quantum Logic*; Springer Science & Business Media: New York, NY, USA, 1998.
41. Engesser, K.; Gabbay, D.M.; Lehmann, D. *Handbook of Quantum Logic and Quantum Structures: Quantum Structures*; Elsevier: Amsterdam, The Netherlands, 2011.
42. Jauch, J.M.; Morrow, R.A. Foundations of quantum mechanics. *Am. J. Phys.* **1968**, *36*, 771. [CrossRef]
43. Piron, C. On the foundations of quantum physics. In *Quantum Mechanics, Determinism, Causality, and Particles*; Springer: Dordrecht, The Netherlands, 1976; pp. 105–116.
44. Kalmbach, G. *Orthomodular Lattices*; Academic Press: London, UK, 1983; Number 18.
45. Kalmbach, G. *Measures and Hilbert Lattices*; World Scientific: Singapore, 1986.
46. Greechie, J. A non-standard quantum logic with a strong set of states. In *Current Issues in Quantum Logic*; Springer: Boston, MA, USA, 1981; pp. 375–380.
47. Giuntini, R. *Quantum Logic and Hidden Variables*; Bibliographisches Institut: Mannheim, Germany, 1991.
48. Pták, P.; Pulmannová, S. *Orthomodular Structures as Quantum Logics*; INIS-mf14676; Kluwer: Dordrecht, The Netherlands, 1991; p. 27.
49. Dvurecenskij, A.; Pulmannová, S. *New Trends in Quantum Structures*; Springer Science & Business Media: New York, NY, USA, 2013; Volume 516.
50. Aerts, D.; Daubechies, I. A characterization of subsystems in physics. *Lett. Math. Phys.* **1979**, *3*, 11–17. [CrossRef]
51. Aerts, D. Classical theories and nonclassical theories as special cases of a more general theory. *J. Math. Phys.* **1983**, *24*, 2441–2453. [CrossRef]

52. Aerts, D. Construction of a structure which enables to describe the joint system of a classical system and a quantum system. *Rep. Math. Phys.* **1984**, *20*, 117–129. [CrossRef]
53. Aerts, D. Construction of the tensor product for the lattices of properties of physical entities. *J. Math. Phys.* **1984**, *25*, 1434–1441. [CrossRef]
54. Beltrametti, E.; Bugajski, S.; Varadarajan, V. Extensions of convexity models. *J. Math. Phys.* **2000**, *41*, 2500–2514. [CrossRef]
55. Barnum, H.; Wilce, A. Information processing in convex operational theories. *Electron. Notes Theor. Comput. Sci.* **2011**, *270*, 3–15. [CrossRef]
56. Barnum, H.; Duncan, R.; Wilce, A. Symmetry, compact closure and dagger compactness for categories of convex operational models. *J. Philos. Log.* **2013**, *42*, 501–523. [CrossRef]
57. Barnum, H.; Barrett, J.; Leifer, M.; Wilce, A. Generalized no-broadcasting theorem. *Phys. Rev. Lett.* **2007**, *99*, 240501. [CrossRef]
58. Barnum, H.; Dahlsten, O.C.; Leifer, M.; Toner, B. Nonclassicality without entanglement enables bit commitment. In Proceedings of the 2008 IEEE Information Theory Workshop, Porto, Portugal, 5–9 May 2008; pp. 386–390.
59. Barnum, H.; Barrett, J.; Clark, L.O.; Leifer, M.; Spekkens, R.; Stepanik, N.; Wilce, A.; Wilke, R. Entropy and information causality in general probabilistic theories. *New J. Phys.* **2010**, *12*, 033024. [CrossRef]
60. Short, A.J.; Wehner, S. Entropy in general physical theories. *New J. Phys.* **2010**, *12*, 033023. [CrossRef]
61. Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **2007**, *75*, 032304. [CrossRef]
62. Perinotti, P. Discord and nonclassicality in probabilistic theories. *Phys. Rev. Lett.* **2012**, *108*, 120502. [CrossRef] [PubMed]
63. Barnum, H.; Barrett, J.; Leifer, M.; Wilce, A. Teleportation in general probabilistic theories. In *Proceedings of Symposia in Applied Mathematics*; American Mathematical Society: Providence, RI, USA, 2012; Volume 71, pp. 25–48.
64. Holik, F.; Massri, C.; Plastino, A. Generalizing entanglement via informational invariance for arbitrary statistical theories. *SOP Trans. Theor. Phys.* **2014**, *1*, 138–153. [CrossRef]
65. Mielnik, B. Geometry of quantum states. *Commun. Math. Phys.* **1968**, *9*, 55–80. [CrossRef]
66. Mielnik, B. Theory of filters. *Commun. Math. Phys.* **1969**, *15*, 1–46. [CrossRef]
67. Mielnik, B. Generalized quantum mechanics. *Commun. Math. Phys.* **1974**, *37*, 221–256. [CrossRef]
68. Khrennikov, A.Y. *Ubiquitous Quantum Structure: From Psychology to Finance*; Springer: Berlin/Heidelberg, Germany, 2010.
69. Narens, L.E. *Probabilistic Lattices: With Applications to Psychology*; World Scientific: Singapore, 2014; Volume 5.
70. Narens, L. Alternative Probability theories for cognitive psychology. *Top. Cogn. Sci.* **2014**, *6*, 114–120. [CrossRef]
71. Rédei, M. The Birkhoff-von Neumann concept of quantum logic. In *Quantum Logic in Algebraic Approach*; Springer: Dordrecht, The Netherlands, 1998; pp. 103–117.
72. Ballentine, L. The statistical interpretation of quantum mechanics. In *Reviews of Modern Physics*; American Physical Society: New York, USA, 1970; Volume 42.
73. Reed, M.; Simon, B. *Methods of Modern Mathematical Physics: Functional Analysis*; Revised Edition; Academic Press: New York, NY, USA, 1980.
74. Laplace, P.S. *Essai Philosophique sur les Probabilités*; H. Remy: Brussels, Belgium, 1829.
75. Jeffreys, H. *Theory of Probability*, 3rd ed.; Clarendon Press: Oxford, UK, 2003.
76. Boole, G. *An Investigation of the Laws of Thought: On Which Are Founded the Mathematical Theories of Logic and Probabilities*; Walton and Maberly: London, UK, 1854; Volume 2.
77. Jaynes, E.T. Information theory and statistical mechanics. II. *Phys. Rev.* **1957**, *108*, 171. [CrossRef]
78. Van Horn, K.S. Constructing a logic of plausible inference: A guide to Cox's theorem. *Int. J. Approx. Reason.* **2003**, *34*, 3–24. [CrossRef]
79. Terenin, A.; Draper, D. Rigorizing and extending the Cox-Jaynes derivation of probability: Implications for statistical practice. *arXiv* **2015**, arXiv:1507.06597.
80. Halpern, J.Y. A counter example to theorems of Cox and Fine. *J. Artif. Intell. Res.* **1999**, *10*, 67–85. [CrossRef]
81. Jaynes, E.T. The well-posed problem. *Found. Phys.* **1973**, *3*, 477–492. [CrossRef]
82. Park, S.Y.; Bera, A.K. Maximum entropy autoregressive conditional heteroskedasticity model. *J. Econom.* **2009**, *150*, 219–230. [CrossRef]
83. Murray, F.J.; von Neumann, J. On rings of operators. *Ann. Math.* **1936**, *37*, 116–229. [CrossRef]
84. Murray, F.J.; Von Neumann, J. On rings of operators. II. *Trans. Am. Math. Soc.* **1937**, *41*, 208–248. [CrossRef]
85. Von Neumann, J. On rings of operators. III. *Ann. Math.* **1940**, *41*, 94–161. [CrossRef]
86. Murray, F.J.; von Neumann, J. On rings of operators. IV. *Ann. Math.* **1943**, *44*, 716–808. [CrossRef]
87. Stubbe, I.; Van Steirteghem, B. Propositional systems, Hilbert lattices and generalized Hilbert spaces. In *Handbook of Quantum Logic and Quantum Structures*; Elsevier: Amsterdam, The Netherlands, 2007; pp. 477–523.
88. Soler, M.P. Characterization of Hilbert spaces by orthomodular spaces. *Commun. Algebra* **1995**, *23*, 219–243. [CrossRef]
89. Gleason, A.M. Measures on the closed subspaces of a Hilbert space. In *The Logico-Algebraic Approach to Quantum Mechanics*; Springer: Dordrecht, The Netherlands, 1975; pp. 123–133.
90. Buhagiar, D.; Chetcuti, E.; Dvurečenskij, A. On Gleason's theorem without Gleason. *Found. Phys.* **2009**, *39*, 550–558. [CrossRef]
91. Bengtsson, I.; Życzkowski, K. *Geometry of Quantum States: An Introduction to Quantum Entanglement*; Cambridge University Press: Cambridge, UK, 2017.

92. Holik, F.; Massri, C.; Plastino, A.; Zuberman, L. On the lattice structure of probability spaces in quantum mechanics. *Int. J. Theor. Phys.* **2013**, *52*, 1836–1876. [CrossRef]

93. Holik, F.; Plastino, A. Convex polytopes and quantum separability. *Phys. Rev. A* **2011**, *84*, 062327. [CrossRef]

94. Bosyk, G.M.; Portesi, M.; Holik, F.; Plastino, A. On the connection between complementarity and uncertainty principles in the Mach–Zehnder interferometric setting. *Phys. Scr.* **2013**, *87*, 065002. [CrossRef]

95. Svozil, K. Classical predictions for intertwined quantum observables are contingent and thus inconclusive. *Quantum Reports* **2020**, *2*, 278–292. [CrossRef]

96. Kochen, S.; Specker, E.P. The problem of hidden variables in quantum mechanics. In *The Logico-Algebraic Approach to Quantum Mechanics*; Springer: Dordrecht, The Netherlands, 1975; pp. 293–328.

97. Cabello, A.; Estebaranz, J.; García-Alcaine, G. Bell-Kochen–Specker theorem: A proof with 18 vectors. *Phys. Lett. A* **1996**, *212*, 183–187. [CrossRef]

98. Jorge, J.P.; Holik, F. Non-deterministic semantics for quantum states. *Entropy* **2020**, *22*, 156. [CrossRef]

99. Cattaneo, G.; Gudder, S. Algebraic structures arising in axiomatic unsharp quantum physics. *Found. Phys.* **1999**, *29*, 1607–1637. [CrossRef]

100. Lahti, P.J.; Busch, P.; Mittelstaedt, P. Some important classes of quantum measurements and their information gain. *J. Math. Phys.* **1991**, *32*, 2770–2775. [CrossRef]

101. Foulis, D.J.; Bennett, M.K. Effect algebras and unsharp quantum logics. *Found. Phys.* **1994**, *24*, 1331–1352. [CrossRef]

102. Foulis, D.J.; Gudder, S.P. Observables, calibration, and effect algebras. *Found. Phys.* **2001**, *31*, 1515–1544. [CrossRef]

103. Busch, P.; Kiukas, J.; Lahti, P. On the notion of coexistence in quantum mechanics. *Math. Slovaca* **2010**, *60*, 665–680. [CrossRef]

104. Heinonen, T. *Imprecise Measurements in Quantum Mechanics*; Turun Yliopisto: Turku, Finland, 2005.

105. Ma, Z.; Zhu, S. Topologies on quantum effects. *arXiv* **2008**, arXiv:0811.2454.

106. Ali, S.T.; Carmeli, C.; Heinosaari, T.; Toigo, A. Commutative POVMs and fuzzy observables. *Found. Phys.* **2009**, *39*, 593–612. [CrossRef]

107. Navara, M.; Rogalewicz, V. The pasting constructions for orthomodular posets. *Math. Nachrichten* **1991**, *154*, 157–168. [CrossRef]

108. Holik, F.; Bosyk, G.M.; Bellomo, G. Quantum information as a non-Kolmogorovian generalization of Shannon's theory. *Entropy* **2015**, *17*, 7349–7373. [CrossRef]

109. Werner, R.F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **1989**, *40*, 4277. [CrossRef] [PubMed]

110. Schrödinger, E. Discussion of probability relations between separated systems. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1935; Volume 31, pp. 555–563.

111. Holik, F.; Massri, C.; Plastino, A. Geometric probability theory and Jaynes's methodology. *Int. J. Geom. Methods Mod. Phys.* **2016**, *13*, 1650025. [CrossRef]

112. Holik, F.; Plastino, A. Quantal effects and MaxEnt. *J. Math. Phys.* **2012**, *53*, 073301. [CrossRef]

113. Aczél, J. *Lectures on Functional Equations and Their Applications*; Academic Press: New York, NY, USA, 1966.