Kolmogorov-type conditional probabilities among distinct contexts

Karl Svozil*
Institute for Theoretical Physics, Vienna University of Technology, Wiedner Hauptstrasse 8-10/136, 1040 Vienna, Austria
(Dated: March 26, 2019)

We suggest applying Kolmogorov’s axioms of probability theory to conditional probabilities among distinct (but not necessarily disjoint and non-intertwining) contexts. Formally, this amounts to row stochastic matrices whose entries characterize the conditional probability to find some observable in one context, given an observable in another context. As the respective probabilities need not (but, depending on the physical/model realization, can be) of the Born rule type, this generalizes approaches to quantum probabilities by Auffèves and Grangier, which in turn are inspired by Gleason’s theorem.

Keywords: Value indefiniteness, Kolmogorov Axioms of probability theory, Pitowsky’s Logical Indeterminacy Principle, Quantum mechanics, Gleason theorem, Kochen-Specker theorem, Born rule

I. QUANTUM BISTOCHASTICITY

In what follows any “largest” domain of mutually commuting observables will be termed context. For quantum mechanics grounded in Hilbert space, a context can be equivalently represented by (i) an orthonormal basis, (ii) the respective one-dimensional orthogonal projection operators associated with the basis elements, or (iii) a single maximal operator whose spectral sum is non-degenerated [1, 2].

An essential assumption entering Gleason’s derivation [3] of the Born rule for quantum probabilities is the validity of classical probability theory whenever the respective observables are co-measurable. Formally, this amounts to the validity of Kolmogorov probability theory for mutually commuting observables; and in particular, to the assumption of Kolmogorov’s axioms within contexts.

Already Gleason pointed out [3] that it is quite straightforward to find an ad hoc conditional probability satisfying this aforementioned assumption, which is based on the Pythagorean property: suppose a pure state formalized by a unit vector and some “measurement frame” formalized by an orthonormal basis. Then the conditional probabilities for the vectors of the orthonormal basis \( \mathcal{E} = \{ |e_1\rangle, \ldots, |e_n\rangle \} \) representing this frame – aka their respective orthogonal projection operators associated with observable propositions – given that pure state \( |\psi\rangle \) can be obtained by taking the absolute square of their scalar products \( \langle |\psi\rangle | e_i \rangle^2 \): Since the vector associated with the pure state, as well as all the vectors in the orthonormal system are of length one, and since these latter vectors (of the orthonormal system) are mutually orthogonal, the sum \( \sum_{i=1}^{n} \langle |\psi\rangle | e_i \rangle^2 \) of all these terms, taken over all the basis elements, needs to add up to one. The respective absolute squares are bounded between zero and one. In effect, the orthonormal basis “grants a view” of the pure quantum state. The absolute square can be rewritten in terms of a trace (over some arbitrary orthonormal basis) into the standard form known as the Born rule of quantum probabilities: \( \langle |\psi\rangle | e_i \rangle^2 = \langle |\psi\rangle | e_i \rangle \langle e_i | \psi \rangle = \langle |\psi\rangle | e_i \rangle \langle e_i | \psi \rangle = \sum_{j=1}^{n} \langle |\psi\rangle | e_i \rangle \langle e_i | |\psi\rangle \rangle (|g_j \rangle \langle g_j | |\psi\rangle \rangle) = \langle |\psi\rangle | e_i \rangle \langle e_i | \psi \rangle = \sum_{j=1}^{n} \langle |\psi\rangle | e_i \rangle \langle e_i | g_j \rangle \langle g_j |\psi \rangle \rangle \)

are the orthogonal projection operators representing the state \( |\psi\rangle \) and \( e_i \), respectively, and \( \mathcal{E}' = \{ |e_1\rangle, \ldots, |e_n\rangle \} \) is an arbitrary orthonormal basis, so that \( I_n = \sum_{j=1}^{n} |g_j \rangle \langle g_j | \rangle \).

It is also well known that, at least from a formal perspective, unit vectors in quantum mechanics serve a dual role: On the one hand, they represent pure states. On the other hand, by the associated one-dimensional orthogonal projection operator, they represent an observable: the proposition that the system is in such a pure state [4, 5]. Suppose now that we exploit this dual role by expanding the pure prepared state into a full orthonormal basis, of which its vector must be an element. (Such an expansion will not be unique as there is a continuous infinity of ways to achieve this.) Once the latter basis is fixed it can be used to obtain a “view” on the former (measurement basis); and a completely symmetric situation/configuration is attained. We might even go so far as to say that which basis is associated with the “observed object” and with the “measurement apparatus,” respectively, is purely a matter of convention, and thus epistemic.

Formally, an orthogonal projection operator serves a dual role: on the one hand it is a formalization of a dichotomic observable – more precisely, an elementary yes-no proposition \( E = |x\rangle \langle x| \) associated with the claim that “the quantized system is in state \( |x\rangle \).” And on the other hand it is the formal representation of a pure quantum state \( |y\rangle \), equivalent to the operator \( F = |y\rangle \langle y| \). By the Born rule the conditional probabilities are symmetric with respect to exchange of \( |x\rangle \) and \( |y\rangle \); let \( \mathcal{E}' = \{ |g_1\rangle, \ldots, |g_n\rangle \} \) be some arbitrary orthonormal basis of \( \mathbb{C}^n \), then \( P(E|F) = \text{Trace}(FE) = \text{Trace}(EF) = P(F|E) \); or, more explicitly, \( P(E|F) = \sum_{j=1}^{n} |g_j \rangle \langle x | |g_j \rangle \langle y | g_j \rangle = \sum_{j=1}^{n} \langle x | y | g_j \rangle \langle g_j | x \rangle = |x \rangle \langle x | y \rangle = P(F|E) \).

Therefore, the respective conditional probabilities form a doubly stochastic (biostochastic) square matrix. This result is a special case of a more general result on quadratic forms on the set of eigenvectors of normal operators [6].

These two orthonormal bases are then two contexts whose respective conditional probabilities can be arranged into a matrix form: its \( r \)th row corresponds to all conditional probabilities associated with the occurrence of the observables of the

* svozil@tuwien.ac.at; http://hph.tuwien.ac.at/~svozil

arXiv:1903.10424v1 [quant-ph] 25 Mar 2019
second context, given the \( i \)th element of the first context. Conversely, its \( j \)th column correspond to all conditional probabilities associated with the occurrence of the observables of the first context, given the \( j \)th element of the second context. By Gleason’s assumption of the validity of Kolmogorov’s axioms within contexts, this matrix of needs to be doubly stochastic (bistochastic) [7, 8]; that is, the sum is taken within every single row and every single column adds up to one.

II. GENERALIZATION OF KOLMOGOROV AXIOMS TO ARBITRARY EVENT STRUCTURES

In order to generalize the quantum case, we suggest to postulate that the quantum case is just one instance satisfying a very general axiom: That, given two arbitrary contexts \( \mathcal{C}_1 = \{e_1, \ldots, e_n\} \) and \( \mathcal{C}_2 = \{f_1, \ldots, f_m\} \), the associated \((n \times m)\)-matrix whose entries are the conditional probabilities \( P(f_j | e_i) \) of “\( f_j \) given \( e_i \)” must be such that the sum taken within every single row adds up to one.

We shall be mostly concerned with cases for which \( n = m \); that is, the associated matrix is a row (aka right) stochastic (square) matrix. Formally, such a matrix \( \mathbf{A} \) has nonnegative entries \( a_{ij} \geq 0 \) for \( i, j = 1, \ldots, n \) whose row sums add up to one: \( \sum_{j=1}^{n} a_{ij} = 1 \) for \( i = 1, \ldots, n \). If, in addition to the row sums, also the column sums add up to one – that is, if \( \sum_{i=1}^{n} a_{ij} = 1 \) for \( j = 1, \ldots, n \) – then the matrix is called doubly stochastic. If \( \mathbf{J} \) is a \((n \times n)\)-matrix whose entries are 1, then a \((n \times n)\)-matrix \( \mathbf{A} \) is row stochastic if \( \mathbf{A} \mathbf{J} = \mathbf{J} \).

The above criterion is a generalization of Kolmogorov’s axioms, as it allows cases in which both contexts do not coincide. For coinciding contexts this rule just reduces to Kolmogorov’s axioms.

III. EXAMPLES OF APPLICATION OF THE GENERALIZED KOLMOGOROV AXIOMS

A. Quantum logics

I have already mentioned the quantum case, which has been studied in an attempt to motivate the Born rule [7, 8]. Therefore, I shall relegate the reader to those studies and proceed to quasi-classical propositional structures and two “exotic” cases.

B. Quasi-classical partition logics

In what follows we shall study sets of partitions of a given set. They have models [9] based on the finite automata initial state identification problem [10] as well as on generalized urns [11, 12]. Partition logics are quasi-classical and value-definite in so far as they allow a separating set of “classical” two-valued states [2, Theorem 0]; and yet they feature complementarity. Many of these logics are doubles of quantum logics, such as for spin-state measurements; and thereby their graphs also allow faithful orthogonal representations [13]; and yet some of them have no quantum analog. Therefore, they neither form a proper subset of all quantum logics, nor do they contain all logical structures encountered in quantum logics (they are neither continuous, nor can they have a nonseparating or nonexisting set of two-valued states). However, those categories overlap significantly, as they bear strong similarities with the structures arising in quantum theory.

If some (partition) logic which is a pasting [14–16] of contexts has a separating set of two-valued states [2, Theorem 0] then there is a constructive, algorithmic [17] way of finding a “canonical” partition logic [9], and, associated with it, all classical probabilities on it: first, find all the two-valued states on the logic, and number these states consecutively. Then, for any atom (element of a context), find the index set of all two-valued states which are 1 on this atom. Associate with each one, say, the \( i \)th of the two valued states a non-negative weight \( i \to \lambda_i \), and require that the (convex) sum of these weights \( \sum \lambda_i = 1 \). Since all two-valued states are included, the Kolmogorov axioms guarantee that the sum of measures/weights within each of the contexts in the logic exactly adds up to one.

Note that in this case, and unlike for quantum conditional probabilities, the conditional probabilities, in general, do not form a bistochastic matrix.

C. Two non-intertwining two-atomic contexts

In the Babylonian spirit [18, p. 172] consider some anecdotal examples which have quantum doubles. The first one will be a spin-\( \frac{1}{2} \) state measurement analogy.

\[
\{1,2\} \quad \{3,4\}
\]

\[
\{1,3\} \quad \{2,4\}
\]

FIG. 1. Greechie orthogonality diagram of a logic with the associated (quasi)classical partition logic representations obtained through in inverse construction using all two-valued measures thereon [9].

The logic in Fig. 1 labels the atoms (aka elementary propositions) obtained by an “inverse construction” using all two-valued measures thereon [9]. With the identifications \( e_1 \equiv \{1,2\}, e_2 \equiv \{3,4\}, f_1 \equiv \{1,3\}, \) and \( f_2 \equiv \{2,4\} \) we obtain all classical probabilities by identifying \( i \to \lambda_i > 0 \). The respec-
tive conditional probabilities are

$$P(\mathcal{G}_2 | \mathcal{G}_1) = [P(\{f_1, f_2\} | \{e_1, e_2\}) $$

$$= \frac{P(f_1 | e_1) P(f_2 | e_1)}{P(f_1, f_2 | e_1)} = \frac{P(f_1 | e_2) P(f_2 | e_2)}{P(f_1, f_2 | e_2)}$$

$$= \frac{P(1, 3) \cap (1, 2)}{P(1, 3) \cap (3, 4)} \frac{P(2, 4) \cap (1, 2)}{P(2, 4) \cap (3, 4)}$$

$$= \frac{P(1)}{P(1, 3)} = \frac{P(2)}{P(2, 4)}$$

$$= \frac{P(3)}{P(3, 4)} \frac{P(4)}{P(4, 5)}$$

$$\lambda_1 \lambda_2 \lambda_3 \lambda_4$$

(1)

as well as

$$P(\mathcal{G}_1 | \mathcal{G}_2) = [P(\{e_1, e_2\} | \{f_1, f_2\}) $$

$$= \frac{P(e_1 | f_1) P(e_2 | f_1)}{P(e_1, e_2 | f_1)} = \frac{P(e_1 | f_2) P(e_2 | f_2)}{P(e_1, e_2 | f_2)}$$

$$= \frac{P(1)}{P(1, 3)} = \frac{P(3)}{P(3, 4)} \frac{P(4)}{P(4, 5)}$$

$$\lambda_1 \lambda_2 \lambda_3 \lambda_4$$

(2)

D. Two intertwining three-atomic contexts

![Diagram of intertwined contexts](image)

FIG. 2. Greechie orthogonality diagram of the L_{12} “firefly” logic with the associated (quasi)classical partition logic representations obtained through in inverse construction using all two-valued measures thereon [9].

The L_{12} “firefly” logic in Fig. 2 labels the atoms (aka elementary propositions) obtained by an “inverse construction” using all two-valued measures thereon [9]. By design, it will be very similar to the earlier logic with four atoms. With the identifications e_1 \equiv \{1, 2\}, e_2 \equiv \{3, 4\}, e_3 \equiv \{5\}, f_1 \equiv \{1, 3\}, and f_2 \equiv \{2, 4\} we obtain all classical probabilities by identifying \(i \rightarrow \lambda_i \). The respective conditional probabilities are

$$P(\mathcal{G}_2 | \mathcal{G}_1) = [P(\{f_1, f_2\} | \{e_1, e_2, e_3\}) $$

$$= \frac{P(1)}{P(1, 3)} = \frac{P(2)}{P(2, 4)}$$

$$\lambda_2 \lambda_3 \lambda_4$$

as well as

$$P(\mathcal{G}_1 | \mathcal{G}_2) = [P(\{e_1, e_2, e_3\} | \{f_1, f_2, f_3\}) $$

$$= \frac{P(1)}{P(1, 3)} = \frac{P(2)}{P(2, 4)}$$

$$\lambda_1 \lambda_2 \lambda_3 \lambda_4$$

(3)

E. Pentagon/pentagram/house logic with five cyclically intertwining three-atomic contexts

By now it should be clear how classical conditional probabilities work on partition logics. Consider the pentagon/pentagram/(orthomodular) house [15, p. 46, Fig. 4.4] logic in Fig. 3 labels the atoms (aka elementary propositions) obtained by an “inverse construction” using all 11 two-valued measures thereon [11]. Take, for example, one of the two contexts $\mathcal{G}_4 = \{\{2, 7, 8\}, \{1, 3, 9, 10, 11\}, \{4, 5, 6\}\}$ “opposite” to the context $\mathcal{G}_1 = \{\{1, 2, 3\}, \{4, 5, 7, 9, 11\}, \{6, 8, 10\}\}$.

![Diagram of pentagon/pentagram/house logic](image)

FIG. 3. Greechie orthogonality diagrams of the pentagon/pentagram/house logic.

With the identifications e_1 \equiv \{1, 2, 3\}, e_2 \equiv \{4, 5, 7, 9, 11\}, e_3 \equiv \{6, 8, 10\}, f_1 \equiv \{2, 7, 8\}, f_2 \equiv \{1, 3, 9, 10, 11\}, and f_3 \equiv \{4, 5, 6\}. The respective conditional probabilities are...
IV. WRIGHT’S TWELFTH DISPERSIONLESS STATE ON THE PENTAGON/PENTAGRAM/HOUSE LOGIC

Despite the aforementioned 11 two-valued states there exists another dispersionless state on cyclic pastings of an odd number of contexts; namely, a state being equal to another dispersionless state on cyclic pastings of an odd number of contexts. The resulting conditional probabilities are neither realizable by quantum nor by classical probability distributions. In this case the conditional probabilities of any two contexts $E_i$ and $E_j$, for $1 \leq i, j \leq 5$ are

$$[P(E_i | E)] = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}. \quad (6)$$

V. THREE-COLORABLE DENSE POINTS ON THE SPHERE

There exist dense subsets of the unit sphere in three dimensions which require just three colors for associating different colors within every mutually orthogonal triple of (unit) vectors \([19-21]\) forming an orthonormal basis. By identifying two of these colors with the value “0”, and the remaining color with the value “1” one obtains a two-valued measure on this “reduced” sphere. The resulting conditional probabilities are discontinuous.

VI. EXTREMA OF CONDITIONAL PROBABILITIES IN ROW AND DOUBLY STOCHASTIC MATRICES

The row stochastic matrices representing conditional probabilities form a polytope in $\mathbb{R}^{n^2}$ whose vertices are the $n^2$ matrices $T_i$, $i = 1, \ldots, n^2$, with exactly one entry 1 in each row \([22, p. 49]\). Therefore, a row stochastic matrix can be represented as the convex sum $\sum_{k=1}^{m} \lambda_k T_i$, with nonnegative $\lambda_k \geq 0$ and $\sum_{k=1}^{m} \lambda_k = 1$.

For conditional probabilities yielding doubly stochastic matrices, such as, for instance, the quantum case, the Birkhoff theorem \([6]\) yields more restricted linear bounds: it states that any doubly stochastic $(n \times n)$-matrix is the convex hull of $m \leq (n - 1)^2 + 1 \leq n!$ permutation matrices. That is, if $A = a_{ij}$ is a doubly stochastic matrix such that $a_{ij} \geq 0$ and $\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n} a_{ij} = 1$ for $1 \leq i, j \leq n$, then there exists a convex sum decomposition $A = \sum_{k=1}^{m} \lambda_k P_k$ in terms of $m \leq (n - 1)^2 + 1 \leq n!$ linear independent permutation matrices $P_k$ such that $\lambda_k \geq 0$ and $\sum_{k=1}^{m} \lambda_k = 1$.

VII. SUMMARY

I have attempted to sketch a generalized probability theory for configurations of observables which have no classical event structure. In particular, if complementarity and distinct contexts are involved this needs an extension of the Kolmogorov axioms. This has been achieved by the requirement that the conditional probabilities of observables in one context, given the occurrence of observables in another context, forms a stochastic matrix.

Various models have been discussed. In the case of doubly stochastic matrices, linear bounds have been derived from the convex hull of permutation matrices.

ACKNOWLEDGMENTS

I kindly acknowledge an invitation to the Santiago based IFICC, as well as enlightening discussions with Tomas Veloz and Philippe Grangier. All misconceptions and errors are mine.

[1] Paul Richard Halmos, *Finite-Dimensional Vector Spaces*, Undergraduate Texts in Mathematics (Springer, New York, 1958).

[2] Simon Kochen and Ernst P. Specker, “The problem of hidden
variables in quantum mechanics,” *Journal of Mathematics and Mechanics* (now Indiana University Mathematics Journal) **17**, 59–87 (1967).

[3] Andrew M. Gleason, “Measures on the closed subspaces of a Hilbert space,” *Journal of Mathematics and Mechanics* (now Indiana University Mathematics Journal) **6**, 885–893 (1957).

[4] John von Neumann, *Mathematische Grundlagen der Quantenmechanik*, 2nd ed. (Springer, Berlin, Heidelberg, 1932, 1996) English translation in Ref. [23].

[5] Garrett Birkhoff and John von Neumann, “The logic of quantum mechanics,” *Annals of Mathematics* **37**, 823–843 (1936).

[6] Marvin Marcus, “Some properties and applications of doubly stochastic matrices,” *The American Mathematical Monthly* **67**, 215–221 (1960).

[7] Alexia Auffèves and Philippe Grangier, “Recovering the quantum formalism from physically realist axioms,” *Scientific Reports* **7**, 43365 (1–9) (2017), arXiv:1610.06164.

[8] Alexia Auffèves and Philippe Grangier, “Extracontextuality and extravalence in quantum mechanics,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **376**, 20170311 (2018), arXiv:1801.01398.

[9] Karl Svozil, “Logical equivalence between generalized urn models and finite automata,” *International Journal of Theoretical Physics* **44**, 745–754 (2005), arXiv:quant-ph/0209136.

[10] Edward F. Moore, “Gedanken-experiments on sequential machines,” in *Automata Studies. (AM-34)*, edited by C. E. Shannon and J. McCarthy (Princeton University Press, Princeton, NJ, 1956) pp. 129–153.

[11] Ron Wright, “The state of the pentagon. A nonclassical example,” in *Mathematical Foundations of Quantum Theory*, edited by A. R. Marlow (Academic Press, New York, 1978) pp. 255–274.

[12] Ron Wright, “Generalized urn models,” *Foundations of Physics* **20**, 881–903 (1990).

[13] Karl Svozil, “Faithful orthogonal representations of graphs from partition logics,” (2018), arXiv:1810.10423.

[14] Richard J. Greechie, “Orthomodular lattices admitting no states,” *Journal of Combinatorial Theory. Series A* **10**, 119–132 (1971).

[15] Gudrun Kalmbach, *Orthomodular Lattices*, London Mathematical Society Monographs, Vol. 18 (Academic Press, London and New York, 1983).

[16] Mirko Navara and Vladimír Rogalewicz, “The pasting constructions for orthomodular posets,” *Mathematische Nachrichten* **154**, 157–168 (1991).

[17] Josef Tkadlec, “2states. A Pascal program,” (2009, 2017), private communication, electronic message from Aug 23nd, 2017.

[18] Otto Neugebauer, *Vorlesungen über die Geschichte der antiken mathematischen Wissenschaften. 1. Band: Vorgriechische Mathematik* (Springer, Berlin, Heidelberg, 1934).

[19] Chris D. Godsil and J. Zaks, “Coloring the sphere,” (1988, 2012), University of Waterloo research report CORR 88-12, arXiv:1201.0486.

[20] David A. Meyer, “Finite precision measurement nullifies the Kochen-Specker theorem,” *Physical Review Letters* **83**, 3751–3754 (1999), arXiv:quant-ph/9905080.

[21] Hans Havlicek, Günther Krenn, Johann Summhammer, and Karl Svozil, “Colouring the rational quantum sphere and the Kochen-Specker theorem,” *Journal of Physics A: Mathematical and General* **34**, 3071–3077 (2001), arXiv:quant-ph/9911040.

[22] Abraham Berman and Robert J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994).

[23] John von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955) German original in Ref. [4].