Contextuality and the fundamental theorems of quantum mechanics

Andreas Döring and ⓘD Markus Frembs

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Journal of
Mathematical Physics

Young Researcher Award

Recognizing the outstanding work of early career researchers

LEARN MORE >>>

AIP Publishing

63, 072103

Contextuality and the fundamental theorems of quantum mechanics

View Online   Export Citation   CrossMark

Andreas Döring[a)] and Markus Frembs[b)] (iD)

AFFILIATIONS

Department of Physics, Imperial College London, Prince Consort Road, London SW7 2AZ, United Kingdom

[a)]Electronic mail: andreas.doering@posteo.de
[b)]Author to whom correspondence should be addressed: markus.frembs13@imperial.ac.uk

ABSTRACT

Contextuality is a key feature of quantum mechanics, as was first brought to light by Bohr [*Albert Einstein: Philosopher-Scientist*, Library of Living Philosophers Vol. VII, edited by P. A. Schilpp (Open Court, 1998), pp. 199–241] and later realized more technically by Kochen and Specker [J. Math. Mech. **17**, 59 (1967)]. Isham and Butterfield put contextuality at the heart of their topos-based formalism and gave a reformulation of the Kochen–Specker theorem in the language of presheaves in Isham and Butterfield [Int. J. Theor. Phys. **37**, 2669 (1998)]. Here, we broaden this perspective considerably (partly drawing on existing, but scattered results) and show that apart from the Kochen–Specker theorem, Wigner's theorem, Gleason's theorem, and Bell's theorem also relate fundamentally to contextuality. We provide reformulations of the theorems using the language of presheaves over contexts and give general versions valid for von Neumann algebras. This shows that a very substantial part of the structure of quantum theory is encoded by contextuality.

I. INTRODUCTION

A. Structural theorems of quantum theory

There are a small number of key theorems in the foundations of quantum theory which throw the differences between classical and quantum in sharp relief. The first and oldest of these is Wigner's theorem[1] from 1931, which shows that each transformation on the set of pure states of a quantum mechanical system that preserves transition probabilities is given by conjugation with a unitary or anti-unitary operator. Hence, the pure state space of a quantum system has a very specific structure.

In 1957, Gleason[2] proved that any assignment of probabilities to projection operators such that probabilities assigned to orthogonal projections add up must already be given by a quantum state. Since projections represent propositions about the values of physical quantities[3] by the spectral theorem, Gleason's result justified the use of the Born rule—30 years after its introduction—when calculating expectation values in quantum mechanics.

Bell[4] proved in 1964 that in local hidden variable theories, there is an upper bound on correlations that can exist between spatially separated subsystems of a composite system. Quantum theory violates this upper bound and hence cannot be (or be replaced by) a local hidden variable theory. There are a number of fine interpretational points, but it is largely accepted today that the violation of Bell's inequality has been confirmed experimentally.

Finally, in 1967, Kochen and Specker[5] showed that under mild and natural conditions, it is mathematically impossible to assign values to all physical quantities simultaneously. Usually, this is phrased as saying that there are no non-contextual value assignments. Since in classical physics, states do assign values to all physical quantities at once, this is a rather strong result and a severe obstacle to any realist interpretation of the quantum formalism.

Each of these landmark theorems singles out a central aspect of quantum theory that distinguishes it from classical physics. The contents of the theorems are very distinct; each one concerns a different structural aspect of quantum theory. Yet, in this article, we will show that, in fact, all these theorems have a common source, which is *contextuality*. This strongly suggests that contextuality is at the heart of quantum theory and is largely responsible for the structural differences between classical and quantum.

B. Contextuality and its mathematical formalization

Contextuality, which was introduced by Bohr,[6,7,75] is a deep concept, and like many other deep concepts in physics, it has taken on a range of meanings and interpretations in the literature.[8] This has led to a certain danger of talking vaguely and at cross-purposes. In order to avoid this, we define precisely what we mean by contextuality and we give a rather minimal and conservative definition that comes with little interpretational baggage.

We say that a physical system has *physical contextuality* if it has some incompatible physical quantities, that is, quantities that cannot be measured simultaneously in an arbitrary state. For all we know today, physical contextuality in this sense is a characteristic feature of all quantum systems, but not of classical systems, and the restrictions on co-measurability are fundamental and not just due to a lack of experimenters' (or theoreticians') finesse.

Even in a physical system with physical contextuality, there are families of compatible, co-measurable physical quantities. (In an extreme case, such a family could consist of a single physical quantity, although this does not occur for quantum systems.) A family of compatible physical quantities is called a *physical context*. We remark the following:

- Our definition of physical contextuality does not refer to actual measurement setups, but we could, as is often done, identify a measurement setup with a physical context, i.e., with the family of physical quantities that is measured by the setup.
- Our definition does not refer to values measured and/or possessed by physical quantities, nor to probabilities. This is not necessary for our purposes, and it avoids many interpretational issues that often cloud the discussion. In particular, we can consider whether non-contextual assignments of values, probabilities, etc., are possible in our kind of contextual theory.
- In order to be compatible according to our definition, two physical quantities must be co-measurable in all states. Hence, two physical quantities that are co-measurable in some states can still be incompatible and lie in different contexts. We are not concerned with weak measurements, weak values, etc.

The definition of physical contextuality is given in a somewhat intuitive manner, since no precise mathematical formalization of "physical quantities" and "states" is provided so far.

We remark—following Kochen and Specker[5] and especially Conway and Kochen[9]—that much less than the full Hilbert space formalism needs to be given: as long as the physical system under consideration (or some subsystem of it) has the physical quantity usually called "spin-1," which can be measured in different directions in space, the system is contextual, since measurements in different directions cannot be performed simultaneously. The mathematical formalism required to describe this situation is a modest part of projective geometry in three dimensions.

Yet, following standard practice, we will assume the usual Hilbert space formalism in which the set of (bounded) physical quantities is mathematically represented by the set $\mathcal{B}(\mathcal{H})_{sa}$ of bounded self-adjoint operators on the Hilbert space $\mathcal{H}$ of the system. The self-adjoint operators form the real part of the complex, noncommutative algebra of bounded operators on the Hilbert space of system. A context is mathematically formalized as a commutative subalgebra of this noncommutative algebra. This can be generalized to von Neumann algebras of physical quantities. We will quickly recall the necessary mathematical background in Sec. II. Moreover, we will introduce the context category and some minimal background on presheaves in order to make this article largely self-contained.

C. Structural theorems and contextuality

We will show that each of the key theorems mentioned above has an *equivalent* reformulation in the language of presheaves over contexts, thus showing the close and sometimes surprising connections between these theorems and contextuality. The prototype of such results is found in the work by Isham, Butterfield, and Hamilton[10,11] on the Kochen–Specker theorem. We will extend their results to the other fundamental theorems by Wigner, Gleason, and Bell and will provide a bigger, more coherent picture of the role of contextuality in the foundations of quantum mechanics.

In Sec. III, Wigner's theorem is treated. The relevant presheaf is trivial, since Wigner's theorem is based on the mere order of contexts, as we will show. The Kochen–Specker theorem and its reformulation are presented in Sec. IV. Here, the so-called spectral presheaf plays the key role. It can be seen as a generalized state space for a quantum system, and the Kochen–Specker theorem is equivalent to the fact that this space has no points in a suitable sense. Gleason's theorem is treated in Sec. V. Its reformulation is based on the so-called probabilistic presheaf, which does have points, i.e., global sections, and these correspond exactly with quantum states. Finally, we consider Bell's theorem and its relations to contextuality in Sec. VI. The relevant presheaf is a bipartite version of the probabilistic presheaf. This presheaf is based on a simple way of composition via contexts, yet it is rich enough to encode all quantum correlations (and not more). In fact, by adding a notion of time orientation in subsystems, it singles out quantum states unambiguously. Section VII concludes this paper.

II. MATHEMATICAL PRELIMINARIES

A. Algebras of physical quantities

Throughout this work, we will take the perspective of *algebraic quantum mechanics*, that is, we emphasize the role of the physical quantities, or observables, and the algebra they form. This means no departure from standard textbook quantum mechanics, just a slightly different perspective that allows for substantial generalizations. We will assume that the physical quantities generate a von Neumann algebra (standard references are, e.g., Refs. 12–14). In the following, we quickly fix notations.

1. Von Neumann algebras

Let $\mathcal{H}$ be the complex Hilbert space of the quantum system under consideration. The algebra of all bounded linear operators on $\mathcal{H}$ is denoted $\mathcal{B}(\mathcal{H})$. If $\mathcal{H}$ is of finite dimension n, then $\mathcal{H} = \mathbb{C}^n$ and $\mathcal{B}(\mathcal{H}) = M_n(\mathbb{C})$, the algebra of all $n \times n$-matrices with complex entries. If $\mathcal{H}$ is infinite-dimensional, then $\mathcal{B}(\mathcal{H})$ carries several interesting topologies (which all coincide in finite dimensions). We will need the weak (operator) topology, the ultraweak topology, and the norm topology. For details on topologies on $\mathcal{B}(\mathcal{H})$ and how they relate to each other, see e.g., Ref. 14.

In particular, $\mathcal{B}(\mathcal{H})$ is closed in the weak topology and hence is a von Neumann algebra. More generally, every weakly closed, unital subalgebra of $\mathcal{B}(\mathcal{H})$ is a von Neumann algebra, and every von Neumann algebra is of this form for a suitable Hilbert space $\mathcal{H}$. The physical quantities of a quantum system are represented by the bounded *self-adjoint operators* in a von Neumann algebra $\mathcal{N}$. The real vector space of self-adjoint operators in $\mathcal{N}$ is denoted $\mathcal{N}_{\text{sa}}$.

By the *spectral theorem*, each self-adjoint operator a is the norm limit of finite real linear combinations of projection operators, i.e., a can be approximated by operators of the form $\sum_{i=1}^n A_i p_i$, where the A_i are real numbers and the p_i are mutually orthogonal projections, that is, $p_i p_j = \delta_{ij} p_i$.

Let S be a family of bounded operators on $\mathcal{H}$. The commutant of S, denoted S', is the set of all operators in $\mathcal{B}(\mathcal{H})$ that commute with all operators in S,

$$S' = \{ b \in \mathcal{B}(\mathcal{H}) \mid \forall a \in S : [a, b] = 0 \}.$$

Von Neumann's *double commutant theorem* (see, e.g., Ref. 12) shows that a subalgebra $\mathcal{N}$ of $\mathcal{B}(\mathcal{H})$ is weakly closed, i.e., a von Neumann algebra if and only if $\mathcal{N} = \mathcal{N}''$. Here, $\mathcal{N}'' = (\mathcal{N}')'$ is the commutant of the commutant of $\mathcal{N}$.

The projections in a von Neumann algebra $\mathcal{N}$ form a complete orthomodular lattice, denoted $\mathcal{P}(\mathcal{N})$. The least upper bound (*join*) of a family of projections $(p_i)_{i \in I}$ in $\mathcal{P}(\mathcal{N})$ is denoted $\bigvee_{i \in I} p_i$, and the greatest lower bound (*meet*) is denoted $\bigwedge_{i \in I} p_i$. In quantum logic, the projection operators are interpreted as propositions about the values of physical quantities. Let a be a self-adjoint operator representing some physical quantity. For simplicity, assume that $a = \sum_{i=1}^n A_i p_i$. Then, the projection p_i represents the proposition "the physical quantity (represented by) a has the value A_i."

2. States on a von Neumann algebra

A state on a von Neumann algebra is a positive linear functional of norm 1. The states form a convex set $\mathcal{S}(\mathcal{N})$, whose extreme points are called *pure* states. In the case of $\mathcal{N} = \mathcal{B}(\mathcal{H})$, the pure states are the familiar vector states. A state ρ is called *multiplicative* if $\rho(ab) = \rho(a)\rho(b)$ for all $a, b \in \mathcal{N}_{\text{sa}}$. This is a strong condition: non-trivial multiplicative states exist only on commutative von Neumann algebras. For these, they are exactly the pure states. A state ρ is called *normal* if $\rho(a_i) \to \rho(a)$ for every monotone increasing net of operators $a_i \in \mathcal{N}$ with least upper bound a.

Just as in standard quantum mechanics, in algebraic quantum theory, (mathematical) states on a von Neumann algebra are interpreted as physical states of the quantum system, assigning expectation values to physical quantities.

3. Jordan algebras and associativity

Instead of the usual multiplication (by composition) of self-adjoint operators, one can also use another product, defined by

$$\forall a, b \in \mathcal{B}(\mathcal{H})_{\text{sa}} : a \circ b := \frac{1}{2}(ab + ba) = \frac{1}{2}\{a, b\}.$$

This is the *Jordan product*, given by the anti-commutator of a and b up to the conventional factor $\frac{1}{2}$. In contrast to the usual product ab, the Jordan product $a \circ b$ is always self-adjoint, even if a and b do not commute. Hence, there is a *real Jordan algebra* $\mathcal{J}(\mathcal{H})_{\text{sa}} = (\mathcal{B}(\mathcal{H})_{\text{sa}}, \circ)$. Its complexification is $\mathcal{J}(\mathcal{H}) = (\mathcal{B}(\mathcal{H}), \circ)$.

Clearly, the Jordan product is commutative, $a \circ b = b \circ a$ for all $a, b \in \mathcal{B}(\mathcal{H})_{\text{sa}}$. More generally, we can associate a weakly closed Jordan algebra $\mathcal{J}(\mathcal{N}) = (\mathcal{N}, \circ)$ to every von Neumann algebra $\mathcal{N}$ by replacing the generally noncommutative product in $\mathcal{N}$ by the commutative Jordan product. There is a "shadow" of noncommutativity left: $\mathcal{J}(\mathcal{N})$ is associative if and only if $\mathcal{N}$ is commutative.

B. Contexts and the context category

1. Mathematical contexts

We begin with the basic definitions of a (mathematical) context and the partially ordered set of all contexts of a quantum system.

Definition 1. Let $\mathcal{H}$ be the Hilbert space of the quantum system under consideration, and let $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ denote the von Neumann algebra of physical quantities of the system. A context is a commutative von Neumann subalgebra of $\mathcal{N}$.

This definition of a (mathematical) context simply encodes the idea that within a (physical) context, which can be identified with a chosen experimental setup, all the physical quantities are compatible, co-measurable, and, hence, are represented mathematically by commuting self-adjoint operators. We denote contexts as $V, \tilde{V}, \dot{V}, \ldots$.

2. The context category

There is a natural partial order on contexts: some contexts are maximal, that is, one cannot add any further self-adjoint operators to them without destroying commutativity. Other contexts are non-maximal; they are commutative von Neumann subalgebras that are properly contained in a larger context. In fact, if V is a non-maximal context, there are many different maximal contexts containing V.

Hence, any quantum system has many different contexts, with smaller ones contained in larger ones. The key idea for all that follows is that one should not just consider a single context of a quantum system, or a small number, but all of them simultaneously. This idea goes back to Isham[10] and has become a fruitful perspective and powerful tool over the last 20 years. The topos approach to quantum theory and many subsequent developments are based on this idea; for an introduction, see, e.g., Ref. 15. The key definition is as follows:

Definition 2. Let $\mathcal{H}$ be the Hilbert space of the quantum system under consideration, and let $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ denote the von Neumann algebra of physical quantities. The context category of the system is the set of all contexts, i.e., the set of all commutative von Neumann subalgebras of $\mathcal{N}$, equipped with inclusion as partial order. The context category is denoted $\mathcal{V}(\mathcal{N})$. If $\mathcal{N} = \mathcal{B}(\mathcal{H})$, then we will simply write $\mathcal{V}(\mathcal{H})$ [instead of $\mathcal{V}(\mathcal{B}(\mathcal{H}))$].

The name *context category* comes from the fact that every partially ordered set (or poset, for short) can also be regarded as a category.[16] The objects are the elements of the poset, and there is an arrow $a \to b$ if and only if $a \leq b$. Hence, in a poset seen as a category, arrows express the order, so there is at most one arrow between any two objects. We will actually need very little category theory in the following, but we will feel free to use some simple and well-established categorical notions where appropriate. The reader not familiar with category theory can equally well read "context poset" for "context category" throughout. If a context $\tilde{V}$ is contained in another context V, we will write $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$ for the inclusion map. Alternatively, one could just write $\tilde{V} \subset V$.

The following, powerful result is due to Harding and Navara:[17]

Theorem 1. *Let $\mathcal{N}$ be a von Neumann algebra not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ or to $M_2(\mathbb{C})$. Then, the context category $\mathcal{V}(\mathcal{N})$ of $\mathcal{N}$ determines the projection lattice $\mathcal{P}(\mathcal{N})$ as an orthomodular lattice up to isomorphism. Conversely, the projection lattice $\mathcal{P}(\mathcal{N})$ determines the poset $\mathcal{V}(\mathcal{N})$ up to isomorphism.*

In fact, Harding and Navara's proof holds more generally for orthomodular lattices with no maximal Boolean sublattices with only four elements [this is why we exclude the cases $\mathcal{N} = \mathbb{C} \oplus \mathbb{C}$ and $\mathcal{N} = M_2(\mathbb{C})$]. The result shows that the context category, i.e., the set of contexts together with the information of how contexts are contained within each other, encodes exactly the same amount of information as the projection lattice. In this sense, contextuality determines quantum logic and vice versa.

Note that the context category $\mathcal{V}(\mathcal{N})$ is just a poset. Its elements, the contexts $V \subseteq \mathcal{N}$, are just "points" within $\mathcal{V}(\mathcal{N})$, without the inner structure. In particular, from the perspective of $\mathcal{V}(\mathcal{N})$, we do not have access to the commutative von Neumann subalgebras, much less to the operators contained within each context. All the structure of $\mathcal{V}(\mathcal{N})$ lies in the order, that is, in the information of how some contexts are contained within others. This makes the Harding–Navara result quite remarkable, since the mere order structure on contexts determines the full structure of the projection lattice.

A brief remark on two-dimensional systems. While the concept of physical contextuality applies to the two-dimensional case $\mathcal{N} = M_2(\mathbb{C})$ also, the noncontextuality constraints between observables, i.e., the order relations between contexts are trivial in this case: the only order relations involve the trivial context, generated by the identity in $M_2(\mathbb{C})$. In turn, taken out the trivial context, there are no order relations in $\mathcal{V}(M_2(\mathbb{C}))$. Since the order relations are at the heart of physical contextuality and since in this work we argue that the latter is behind many key theorems in quantum foundations, in our reformulations of these theorems, we will exclude the two-dimensional case. One way to justify this is to view any such system as part of a larger system (including an "environment") in which case it immediately embeds within the realm of validity of the theorems as given below.

3. Contexts without (non)commutativity

The context category $\mathcal{V}(\mathcal{N})$ of a von Neumann algebra $\mathcal{N}$ can be defined without any reference to (non)commutativity: every weakly closed associative Jordan subalgebra of $\mathcal{N}$ is a commutative von Neumann subalgebra and vice versa. Hence, we can regard $\mathcal{N}$ as a weakly

closed Jordan algebra and consider the set of its weakly closed associative Jordan subalgebras, partially ordered by inclusion. This poset is (isomorphic to) $\mathcal{V}(\mathcal{N})$.

C. Presheaves over the context category

1. The concept of a presheaf: Local data glued together

We saw in Subsection II B that the context category $\mathcal{V}(\mathcal{N})$ already encodes a lot of information about a quantum system. Now, we will build further structures upon the context category in order to make it an even more useful tool. Concretely, given the context category $\mathcal{V}(\mathcal{N})$, we are interested in assigning data to each context. Moreover, since $\mathcal{V}(\mathcal{N})$ is a poset, we want to relate the data assigned to a context V to the data assigned to a smaller context $\tilde{V} \subset V$.

For example, for each context $V \in \mathcal{V}(\mathcal{N})$, one may consider the set $\Sigma(V)$ of all pure states of V. If $\tilde{V} \subset V$, then every pure state of V gives a pure state of $\tilde{V}$ simply by restriction,

$$\Sigma(V) \longrightarrow \Sigma(\tilde{V}),$$
$$\lambda \longmapsto \lambda|_{\tilde{V}}.$$

(Pure states of commutative von Neumann algebras are traditionally denoted λ.) In this way, we obtain a natural map from the pure states of V to the pure states of $\tilde{V}$, hence relating the data assigned to V to the data assigned to $\tilde{V}$.

This is an example of a general construction, viz., a presheaf. This naming is traditional and has no particular meaning for us. The general definition of a presheaf over the context category $\mathcal{V}(\mathcal{N})$ is as follows:

*Definition 3. Let $\mathcal{N}$ be a von Neumann algebra, and let $\mathcal{V}(\mathcal{N})$ be its context category. A presheaf over $\mathcal{V}(\mathcal{N})$ is a contravariant functor $\underline{P} : \mathcal{V}(\mathcal{N}) \to$ **Set***. That is, $\underline{P}$ is given as follows:*

   (i)   *on objects: for all $V \in \mathcal{V}(\mathcal{N})$, $\underline{P}_V$, the component of $\underline{P}$ at V, is some set;*
   (ii)   *on arrows: for all inclusions $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$, there is a restriction map*

$$\underline{P}(i_{\tilde{V}V}) : \underline{P}_V \longrightarrow \underline{P}_{\tilde{V}},$$
$$x \longmapsto \underline{P}(i_{\tilde{V}V})(x).$$

Of course, this is a very general notion. The idea is that the "local" data assigned to each context can vary from context to context, but there are "connecting maps" relating the local data at V and $\tilde{V}$ whenever $\tilde{V} \subset V$. The mathematical language of presheaves is a convenient tool for book-keeping.

Example 1. The trivial presheaf. The simplest (non-empty) presheaf over $\mathcal{V}(\mathcal{N})$ is the trivial presheaf $\underline{1}$, which is given as follows:

   (i)   *on objects: for all $V \in \mathcal{V}(\mathcal{N})$, $\underline{1}_V := \{*\}$, the one-element set;*
   (ii)   *on arrows: for all inclusions $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$,*

$$\underline{1}(i_{\tilde{V}V}) : \underline{1}_V \longrightarrow \underline{1}_{\tilde{V}},$$
$$* \longmapsto *.$$

We will make use of the trivial presheaf when we consider Wigner's theorem in Sec. III. For the treatment of the Kochen–Specker theorem in Sec. IV, we will use the *spectral presheaf* (already sketched above); for Gleason's theorem (in Sec. V), we will use the *probabilistic presheaf*; and for Bell's theorem (in Sec. VI), we will use the *Bell presheaf*, a version of the probabilistic presheaf adapted to composite systems. Each of these presheaves is tailor-made for reformulating the respective theorem. The spectral presheaf is built from pure states in each context, and the probabilistic presheaf is built from mixed states.

2. Contravariance and coarse-graining

One may wonder why we are using contravariant functors rather than covariant ones. If V, $\tilde{V}$ are contexts such that $\tilde{V} \subset V$, why not map the local data assigned to $\tilde{V}$ into the local data assigned to V? Covariant functors do have their place in the bigger scheme,[18] but as it turns out, for our purposes, we only need contravariant functors. Generically, the idea is that the data assigned to a larger context V is richer, more informative, and can be coarse-grained or restricted to the data assigned to a smaller context $\tilde{V}$. Conversely, there is often no canonical way to "fine-grain" or extend data assigned to $\tilde{V}$ to data assigned to V. It is always possible to discard information, but it is often impossible to create information, at least not in a unique way. For example, every pure state of V gives a pure state of $\tilde{V}$ by restriction, but a pure state of $\tilde{V}$ can usually be extended in many different ways to a pure state of V. (The problem here is that there is no *canonical* extension.)

3. Mapping a presheaf into itself

In order to relate Wigner's theorem to the trivial presheaf in Sec. III, we have to consider automorphisms of the trivial presheaf, that is, reversible mappings of $\underline{1}$ into itself. There are a number of possible definitions and conventions, but we will focus on a very simple notion of automorphism that suits our purposes.

Let $\underline{P}$ be a presheaf over the context category $\mathcal{V}(\mathcal{N})$. Roughly speaking, we can map $\underline{P}$ to itself by first shifting the components around and then mapping each (shifted) component into itself in a way that is compatible with the restriction maps (natural transformation). More precisely, the shifting around of components is achieved by a morphism of the base category, which is $\mathcal{V}(\mathcal{N})$ in our case, acting by pullback. This means that if $\check{\phi} : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ is a morphism of the base category, then it acts as follows: $\underline{P} \circ \check{\phi}$ is the presheaf over $\mathcal{V}(\mathcal{N})$ with component $(\underline{P} \circ \check{\phi})_V = \underline{P}_{\check{\phi}(V)}$ at V. That is, the new component at V is the old component at $\check{\phi}(V)$ for every $V \in \mathcal{V}(\mathcal{N})$. The restriction maps of $\underline{P} \circ \check{\phi}$ are given in the obvious way by $(\underline{P} \circ \check{\phi})(i_{\tilde{V}V}) = \underline{P}(i_{\check{\phi}(\tilde{V})\check{\phi}(V)})$.

An automorphism $\Theta : \underline{P} \to \underline{P}$ of a presheaf $\underline{P}$ over $\mathcal{V}(\mathcal{N})$ then consists of the following:

(a) an automorphism $\check{\phi} : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ of the base category acting by pullback, thus mapping $\underline{P}$ to $\underline{P} \circ \check{\phi}$,
(b) followed by, for each $V \in \mathcal{V}(\mathcal{N})$, an isomorphism $\theta_V : (\underline{P} \circ \check{\phi})_V \to (\underline{P} \circ \check{\phi})_V$ such that, whenever $\tilde{V} \subset V$, one has $\underline{P}(i_{\check{\phi}(\tilde{V})\check{\phi}(V)}) \circ \theta_V = \theta_{\tilde{V}} \circ \underline{P}(i_{\check{\phi}(\tilde{V})\check{\phi}(V)})$.

4. Local and global sections of a presheaf

Presheaves over the context category $\mathcal{V}(\mathcal{H})$ [or $\mathcal{V}(\mathcal{N})$] are not just sets but collections of sets (one for each context), which are interconnected by the restriction maps. Hence, the notion of an "element" of a presheaf must be defined suitably. Let $\underline{P}$ be a presheaf over $\mathcal{V}(\mathcal{N})$, and let D be a downward closed subset of $\mathcal{V}(\mathcal{N})$, i.e., if $V \in D$ and $\tilde{V} \subset V$, then $\tilde{V} \in D$. A *local section* γ of $\underline{P}$ over D consists of a choice of one element from the component $\underline{P}_V$ for each $V \in D$, denoted γ_V, such that whenever $V, \tilde{V} \in D$ with $\tilde{V} \subset V$, then $\underline{P}(i_{\tilde{V}V})(\gamma_V) = \gamma_{\tilde{V}}$. This condition means that the elements γ_V that we pick from the sets $\underline{P}_V$ (where $V \in D$) fit together under the restriction maps of the presheaf $\underline{P}$.

One should think of a local section γ of $\underline{P}$ over D as a "partial" element of $\underline{P}$. If one has a local section over $D = \mathcal{V}(\mathcal{N})$, then γ is called a *global section* (or *global element*) of the presheaf $\underline{P}$. This is the analog of an element of a set or a point of a space.

For a given presheaf $\underline{P}$, global sections may or may not exist (while local sections always exist, just make D small enough). In fact, finding a global section amounts to fitting specified local data, one element from each component of $\underline{P}$, together into a whole. We will see that the presheaf reformulations of the Kochen–Specker theorem (following Isham, Butterfield, and Hamilton), Gleason's theorem, and also Bell's theorem are statements about the existence or nonexistence of global sections of certain presheaves.

III. REFORMULATION OF WIGNER'S THEOREM

A. Wigner's theorem, Dye's theorem, and Jordan *-automorphisms

We first consider the following Wigner's theorem:[1]

Theorem 2 *(Wigner's theorem). Let $\mathcal{H}$ be a Hilbert space, $\dim(\mathcal{H}) \geq 2$, and let $\mathcal{P}_1(\mathcal{H})$ be the set of rank-1 projections on $\mathcal{H}$ [equivalently, $\mathcal{P}_1(\mathcal{H})$ is the projective Hilbert space]. Every bijective map*

$$\varphi : \mathcal{P}_1(\mathcal{H}) \longrightarrow \mathcal{P}_1(\mathcal{H}),$$
$$p \longmapsto \varphi(p)$$

such that $\mathrm{tr}[\varphi(p)\varphi(q)] = \mathrm{tr}[pq]$ *for all* $p, q \in \mathcal{P}_1(\mathcal{H})$ *(i.e., transition probabilities are preserved) is implemented by conjugation with a unitary or anti-unitary operator u,*

$$\forall p \in \mathcal{P}_1(\mathcal{H}) : \varphi(p) = upu^*.$$

Various nice proofs can be found in the literature; for a modern perspective, see, e.g., Ref. 19. These authors also prove the following: let $\mathrm{Aut}(\mathcal{P}_1(\mathcal{H}))$ denote the group of automorphisms of $\mathcal{P}_1(\mathcal{H})$ [i.e., bijective maps $\varphi : \mathcal{P}_1(\mathcal{H}) \to \mathcal{P}_1(\mathcal{H})$ that preserve transition probabilities]. If $\dim(\mathcal{H}) \geq 3$, then $\mathrm{Aut}(\mathcal{P}_1(\mathcal{H}))$ is isomorphic to the group $\mathrm{Aut}(\mathcal{P}(\mathcal{H}))$ of automorphisms of the projection lattice $\mathcal{P}(\mathcal{H})$, i.e., maps

$$\phi : \mathcal{P}(\mathcal{H}) \longrightarrow \mathcal{P}(\mathcal{H}),$$
$$p \longmapsto \phi(p)$$

that

1. are bijective,
2. preserve complements, $\forall p \in \mathcal{P}(\mathcal{H}) : \phi(1 - p) = 1 - \phi(p)$, and

3. preserve and reflect order, $\forall p, q \in \mathcal{P}(\mathcal{H}) : (p \leq q) \Leftrightarrow (\phi(p) \leq \phi(q))$.

Geometrically, $p \leq q$ means that the closed subspace that p projects onto is contained in the closed subspace that q projects onto. Algebraically, $(p \leq q) \Leftrightarrow (pq = p)$. Since an automorphism $\phi \in \mathrm{Aut}(\mathcal{P}(\mathcal{H}))$ preserves the order, it also preserves all meets (greatest lower bounds) and all joins (least upper bounds) in $\mathcal{P}(\mathcal{H})$. Since ϕ also preserves complements, it is an automorphism of the complete orthomodular lattice $\mathcal{P}(\mathcal{H})$.

Under the group isomorphism $\mathrm{Aut}(\mathcal{P}_1(\mathcal{H})) \to \mathrm{Aut}(\mathcal{P}(\mathcal{H}))$, the automorphism $\phi \in \mathrm{Aut}(\mathcal{P}(\mathcal{H}))$ corresponding to a given $\varphi \in \mathrm{Aut}(\mathcal{P}_1(\mathcal{H}))$ is an extension of φ to all projections that preserves joins and complements (and hence also meets; for details, see Ref. 19). Of course, we have $\phi|_{\mathcal{P}_1(\mathcal{H})} = \varphi$.

Hence, if the Hilbert space is at least three-dimensional, Wigner's theorem is equivalent to the fact that every automorphism of the projection lattice $\mathcal{P}(\mathcal{H})$ is implemented by conjugation with a unitary or anti-unitary operator,

$$\forall p \in \mathcal{P}(\mathcal{H}) : \phi(p) = upu^*.$$

There is a generalization of Wigner's theorem to von Neumann algebras, which is closer to the formulation with automorphisms of $\mathcal{P}(\mathcal{H})$ than automorphisms of $\mathcal{P}_1(\mathcal{H})$. This is Dye's theorem.[20] Before we formulate the theorem, we recall that given a von Neumann algebra $\mathcal{N}$, we can form the associated Jordan algebra $\mathcal{J}(\mathcal{N}) = (\mathcal{N}, \circ)$, which has the same elements and linear structure as $\mathcal{N}$, and the Jordan product given by

$$\forall a, b \in \mathcal{N} : a \circ b := \frac{1}{2}(ab + ba) = \frac{1}{2}\{a, b\}.$$

Recall that the Jordan product is not only commutative but also associative only if the von Neumann algebra $\mathcal{N}$ is commutative. A Jordan $*$-automorphism of $\mathcal{J}(\mathcal{N})$ is a bijective map $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{N})$ such that both Φ and Φ^{-1} preserve Jordan product and involution $(\_)^*$,

$$\forall a, b \in \mathcal{J}(\mathcal{N}) : \Phi(a \circ b) = \Phi(a) \circ \Phi(b),$$
$$\forall a \in \mathcal{J}(\mathcal{N}) : \Phi(a^*) = \Phi(a)^*$$

and analogously for Φ^{-1}. We can now formulate Dye's theorem.

Theorem 3. *Let $\mathcal{N}$ be a von Neumann algebra with no direct summand of type I_2. For every automorphism $\phi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{N})$ of the projection lattice of $\mathcal{N}$, there exists a unique Jordan $*$-automorphism $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{N})$ such that $\Phi(p) = \phi(p)$ for all projections $p \in \mathcal{P}(\mathcal{N})$.*

It is easy to see that the Jordan $*$-automorphism Φ induced by an automorphism $\phi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{N})$ is *ultraweakly continuous* (or normal), i.e., it preserves (countable) joins of projections.[21] Conversely, every ultraweakly continuous Jordan $*$-automorphism $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{N})$ induces an automorphism ϕ of the complete orthomodular lattice $\mathcal{P}(\mathcal{N})$ by $\phi := \Phi|_{\mathcal{P}(\mathcal{N})}$.

The ultraweakly continuous Jordan $*$-automorphisms of $\mathcal{J}(\mathcal{N})$ form a group denoted $\mathrm{Aut}(\mathcal{J}(\mathcal{N}))$. Dye's theorem, hence, shows that, provided $\mathcal{N}$ has no type I_2 summand, there is a group isomorphism

$$\mathrm{Aut}(\mathcal{P}(\mathcal{N})) \to \mathrm{Aut}(\mathcal{J}(\mathcal{N}))$$

between the group of automorphisms of the projection lattice and the group of ultraweakly continuous Jordan $*$-automorphisms of $\mathcal{N}$.

One may wonder how Dye's theorem and Jordan $*$-automorphisms relate to unitary and anti-unitary operators as in Wigner's theorem. To see this, we first need the following well-known result (see, e.g., Ref. 22):

Proposition 4. Every Jordan $*$-automorphism $\Phi : \mathcal{N} \to \mathcal{N}$ of a von Neumann algebra $\mathcal{N}$ can be decomposed as the sum of a $*$-isomorphism and a $*$-anti-isomorphism.

More concretely, there are projections p, q in the center of $\mathcal{N}$ such that $\mathcal{N}$ is unitarily equivalent to both $p\mathcal{N} \oplus (1-p)\mathcal{N}$ and $q\mathcal{N} \oplus (1-q)\mathcal{N}$, and $\Phi|_{\mathcal{N}p} : p\mathcal{N} \to q\mathcal{N}$ is a $*$-isomorphism, while $\Phi|_{(1-p)\mathcal{N}} : (1-p)\mathcal{N} \to (1-q)\mathcal{N}$ is a $*$-anti-isomorphism. Moreover, we need the following proposition (see e.g., Ref. 22):

Proposition 5. Every $*$-automorphism $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is implemented by conjugation with a unitary operator and every $*$-anti-automorphism is implemented by conjugation with an anti-unitary operator.

By Dye's theorem and Proposition 4, every (ultraweakly continuous) Jordan */automorphism $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ decomposes into a $*$-isomorphism on $\mathcal{B}(\mathcal{H})p$ and a $*$-anti-isomorphism on $\mathcal{B}(\mathcal{H})(1-p)$, where p is a *central* projection in $\mathcal{P}(\mathcal{H})$. Since $\mathcal{B}(\mathcal{H})$ is a factor (a von Neumann algebra with trivial center), the only central projections are 0 and 1. Thus, a Jordan $*$-automorphism $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is either a $*$-automorphism or a $*$-anti-automorphism. Hence, by Proposition 5, it is of the form

$$\Phi(a) = uau^*$$

for a unitary or anti-unitary operator u acting on $\mathcal{B}(\mathcal{H})$. We see that Wigner's theorem is a special case of Dye's theorem [depending on some special features of $\mathcal{B}(\mathcal{H})$], and we can rephrase Wigner's theorem as follows:

Theorem 6 *(Wigner's theorem in "Jordan formulation"). Let $\mathcal{H}$ be a Hilbert space, $\dim \mathcal{H} \geq 3$. Every automorphism $\phi : \mathcal{P}(\mathcal{H}) \to \mathcal{P}(\mathcal{H})$ is implemented by a unique ultraweakly continuous Jordan $*$-automorphism $\Phi : \mathcal{J}(\mathcal{H}) \to \mathcal{J}(\mathcal{H})$ such that $\phi(p) = \Phi(p)$ for all $p \in \mathcal{P}(\mathcal{H})$.*

This shows that contrary to the usual hand-waving arguments, there is a good mathematical reason to consider both unitary and anti-unitary operators in Wigner's theorem, since we actually have a statement about the structure of $\mathcal{B}(\mathcal{H})$ as a Jordan algebra $\mathcal{J}(\mathcal{H}) = (\mathcal{B}(\mathcal{H}), \circ)$. The Jordan structure is preserved by the action of both unitary and anti-unitary operators. For a related treatment of symmetries in quantum theory, see also Ref. 23.

B. Contextuality and Jordan structure

So far, all this does not relate to contexts in any obvious way. The result that connects Wigner's theorem (and Dye's theorem) with contextuality is the following theorem:

Theorem 7 *(Döring and Harding[24]). Let $\mathcal{N}$ be a von Neumann algebra not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ or to $M_2(\mathbb{C})$. For every order automorphism $\tilde{\phi} : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ of the context category of $\mathcal{N}$, there is a unique (ultraweakly continuous) Jordan $*$-automorphism $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{N})$ such that $\tilde{\phi}(V) = \Phi[V]$ for all $V \in \mathcal{V}(\mathcal{N})$.*

This shows that the mere order structure of contexts determines the algebra of observables as a Jordan algebra up to isomorphism. The proof proceeds in two steps: first, using the result by Harding and Navara[17] already mentioned in Sec. II B, one shows that an order automorphism $\tilde{\phi} : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ induces a unique automorphism $\phi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{N})$ of the projection lattice; second, by Dye's theorem, this gives a Jordan $*$-automorphism $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{N})$. As a shorthand, contextuality determines the Jordan algebra of physical quantities and vice versa.

As remarked elsewhere,[21] it is easy to see that the Jordan $*$-automorphism Φ induced by an order isomorphism $\tilde{\phi} : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ is ultraweakly continuous. Hence, there is a group isomorphism

$$\mathrm{Aut}(\mathcal{V}(\mathcal{N})) \longrightarrow \mathrm{Aut}(\mathcal{J}(\mathcal{N})).$$

The Döring–Harding result can be regarded as a reformulation of Dye's theorem (Theorem 3), explicitly showing that it is contextuality that determines the Jordan algebra structure of von Neumann algebras (and vice versa). Specializing to the algebra $\mathcal{N} = \mathcal{B}(\mathcal{H})$ and using Proposition 5, we have the following theorem:

Theorem 8 *(Wigner's theorem in contextual form). Let $\mathcal{H}$ be a Hilbert space, $\dim(\mathcal{H}) \geq 3$. For every order automorphism $\tilde{\phi} : \mathcal{V}(\mathcal{H}) \to \mathcal{V}(\mathcal{H})$, there is a unique unitary or anti-unitary operator u such that*

$$\forall V \in \mathcal{V}(\mathcal{H}) : \quad \tilde{\phi}(V) = uVu^*.$$

Conversely, every unitary or anti-unitary operator u induces an order automorphism of the context category $\mathcal{V}(\mathcal{H})$ by conjugation.

This is our first reformulation of Wigner's theorem. Remarkably, any bijective map ϕ that preserves the order on the collection of contexts must be implemented by a unitary or anti-unitary operator.

C. The trivial presheaf

Obviously, the structure of the context category $\mathcal{V}(\mathcal{H})$ is sufficient to reformulate Wigner's theorem. Extra information, as may be provided by presheaves over $\mathcal{V}(\mathcal{H})$, is not necessary. In this sense, Wigner's theorem is the simplest of the theorems that we consider.

Yet, there is a reformulation of Wigner's theorem, more generally Dye's theorem, that does use a presheaf. Since we need exactly the information provided by the context category $\mathcal{V}(\mathcal{N})$, which is a partially ordered set, the presheaf must mirror this partial order (and nothing more). The trivial presheaf $\underline{1}$ over $\mathcal{V}(\mathcal{N})$ does this: the component at $V \in \mathcal{V}(\mathcal{N})$ is the one-element set $\{*\}$, and for every inclusion $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$, there is a restriction $\underline{1}(i_{\tilde{V}V}) : \underline{1}_V \to \underline{1}_{\tilde{V}}$, sending $\{*\}$ to $\{*\}$. Hence, we have both the elements of the poset $\mathcal{V}(\mathcal{N})$ and the order relation encoded by $\underline{1}$.[25]

As discussed in Sec. II C, an automorphism Θ of a presheaf $\underline{P}$ consists of two things (according to our convention): a shifting around of components, induced by an automorphism $\tilde{\phi}$ of the base category acting by pullback, followed by an isomorphism $\theta_V : (\underline{P} \circ \tilde{\phi})_V \to (\underline{P} \circ \tilde{\phi})_V$ for each $V \in \mathcal{V}(\mathcal{N})$ such that, whenever $\tilde{V} \subset V$, one has $\underline{P}(i_{\tilde{\phi}(\tilde{V})\tilde{\phi}(V)}) \circ \theta_V = \theta_{\tilde{V}} \circ \underline{P}(i_{\tilde{\phi}(\tilde{V})\tilde{\phi}(V)})$.

In our case, an order automorphism $\tilde{\phi} : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ acts by pullback on the trivial presheaf $\underline{1}$ over $\mathcal{V}(\mathcal{N})$ in the following way: for all $V \in \mathcal{V}(\mathcal{N})$, $(\underline{1} \circ \tilde{\phi})_V = \underline{1}_{\tilde{\phi}(V)} = \{*\}$.

Since each component $\underline{1}_V$ of the trivial presheaf $\underline{1}$ is just a one-element set, $\underline{1}_V = \{*\}$, the only isomorphism $\theta_V : (\underline{1} \circ \tilde{\phi})_V \rightarrow (\underline{1} \circ \tilde{\phi})_V$ is the identity map [for each $V \in \mathcal{V}(\mathcal{N})$]. Hence, an automorphism of the trivial presheaf $\underline{1}$ over $\mathcal{V}(\mathcal{N})$ is simply given by a shifting around of components, induced by an (order) automorphism of the base category $\mathcal{V}(\mathcal{N})$.

Lemma 9. Let $\mathcal{N}$ be a von Neumann algebra. There is a bijective correspondence between automorphisms Θ of the trivial presheaf $\underline{1}$ over $\mathcal{V}(\mathcal{N})$ and order automorphisms $\tilde{\phi}$ of the context category $\mathcal{V}(\mathcal{N})$.

From this and Theorem 7, we have the following corollary:

Corollary 10 (Dye's theorem in presheaf form). Let $\mathcal{N}$ be a von Neumann algebra not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ or to $M_2(\mathbb{C})$. For every automorphism Θ of the trivial presheaf $\underline{1}$ over the context category $\mathcal{V}(\mathcal{N})$, there is a unique (ultraweakly continuous) Jordan $*$-automorphism $\Phi : \mathcal{J}(\mathcal{N}) \rightarrow \mathcal{J}(\mathcal{N})$ such that $\tilde{\phi}(V) = \Phi[V]$ for all $V \in \mathcal{V}(\mathcal{N})$, where $\tilde{\phi} : \mathcal{V}(\mathcal{N}) \rightarrow \mathcal{V}(\mathcal{N})$ is the automorphism of the context category corresponding to Θ.

Finally, from Lemma 9 and Theorem 8 we obtain the following corollary:

Corollary 11 (Wigner's theorem in presheaf form). Let $\mathcal{H}$ be a Hilbert space, $\dim(\mathcal{H}) \geq 3$. For every automorphism Θ of the trivial presheaf $\underline{1}$ over the context category $\mathcal{V}(\mathcal{H})$, there is a unique unitary or anti-unitary operator u such that

$$\forall V \in \mathcal{V}(\mathcal{H}) : \quad \tilde{\phi}(V) = uVu^*$$

is the order automorphism of $\mathcal{V}(\mathcal{H})$ inducing $\Theta : \underline{1} \rightarrow \underline{1}$. Conversely, every unitary or anti-unitary operator u induces an automorphism Θ of the trivial presheaf $\underline{1}$ over $\mathcal{V}(\mathcal{H})$.

Note that the trivial presheaf contains exactly the right amount of information: every automorphism of $\underline{1}$ gives a unitary or anti-unitary u and vice versa. More physically speaking, every "rearrangement" of contexts that preserves the order (i.e., preserves how contexts are contained within each other) determines a unitary or anti-unitary operator and vice versa.

IV. REFORMULATION OF THE KOCHEN–SPECKER THEOREM

The Kochen–Specker theorem[5] is deeply connected to contextuality. The usual interpretation of the theorem amounts to a negative statement of the kind "there are no non-contextual assignments of values to physical quantities."[26]

Kochen and Specker's result excludes certain state space models for quantum theory. This makes it difficult to interpret the Kochen–Specker theorem in geometric terms. In addition, it is not straightforward to see the exact nature of the connection between contextuality in our sense (commutative subalgebras of compatible physical quantities, arranged into a poset) and the nonexistence of valuation functions. Both aspects were clarified by Isham and Butterfield in a beautiful series of papers,[10,11,27,28] with Hamilton as a co-author of the third paper. In fact, the context category first shows up in these papers, and so does the spectral presheaf $\underline{\Sigma}$, which will be defined below. The latter plays a central role in the topos approach to quantum theory[15] and serves as a generalized state space for a quantum system, notwithstanding the Kochen–Specker theorem. In fact, the Kochen–Specker theorem is equivalent to the fact that the quantum state space $\underline{\Sigma}$ has no points (technically, it has no global sections). This reformulation serves as the prototype for the reformulations of the other fundamental theorems of quantum theory discussed in this article.

In Sec. IV A, we will give a quick overview of the Kochen–Specker theorem and its background. Then, in Sec. IV B, we make some connections with contextuality and present the presheaf reformulation of the Kochen–Specker theorem by Isham, Butterfield, and Hamilton. Finally, we extend their results to von Neumann algebras.

A. Valuation functions and the Kochen–Specker theorem

In their seminal paper,[5] Kochen and Specker considered the question whether assignments of values to the physical quantities of a quantum system exist. Let $\mathcal{H}$ be a separable Hilbert space, let $\mathcal{B}(\mathcal{H})$ be the algebra of bounded linear operators on $\mathcal{H}$, and let $\mathcal{B}(\mathcal{H})_{\mathrm{sa}}$ be the real vector space of bounded self-adjoint operators. A *valuation function* is a function

$$v : \mathcal{B}(\mathcal{H})_{\mathrm{sa}} \longrightarrow \mathbb{R},$$
$$a \longmapsto v(a)$$

such that

(i) for all $a \in \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$, it holds that $v(a) \in sp(a)$ (spectrum rule),
(ii) for all continuous functions $f : \mathbb{R} \rightarrow \mathbb{R}$, it holds that $v(f(a)) = f(v(a))$ (*functional composition principle*).

Theorem 12 *(Kochen–Specker theorem[5]). Let $\mathcal{H}$ be a Hilbert space, $\dim \mathcal{H} \geq 3$. There exist no valuation functions $v : \mathcal{B}(\mathcal{H})_{\mathrm{sa}} \rightarrow \mathbb{R}$.*

In the proof, a certain family of rays, i.e., rank-1 projections, is considered. Each of these must be assigned either 0 or 1 according to the spectrum rule and in every orthogonal triple p_1, p_2, p_3 of rank-1 projections, exactly one projection is assigned 1 and the others are assigned 0. By carefully choosing the family of projections, Kochen and Specker constructed an explicit counterexample: they show that no consistent assignment of values 0, 1 to the projections in their family is possible. The original proof used a configuration of 117 rays in $\mathcal{H} = \mathbb{R}^3$ (the real Euclidean space), which could later be reduced to 31 and even fewer in $\mathbb{C}^4$. The proof of the result in real, three-dimensional Hilbert space implies the result in higher-dimensional, real and complex Hilbert spaces.

Bell provided a proof of the same result, i.e., there are no non-contextual assignments of values to physical quantities.[29] His proof uses a continuity argument and Gleason's theorem and hence is not "discrete" as Kochen and Specker's proof.

B. The spectral presheaf

Kochen and Specker emphasized that an earlier proof of nonexistence of certain value assignments by von Neumann[30] was flawed, since it posed conditions on noncompatible physical quantities (represented by noncommuting self-adjoint operators), which Kochen and Specker regarded as unjustified. In contrast, the functional composition principle employed by Kochen and Specker seemingly is just a condition on commuting operators, since a and $f(a)$ commute.

Yet, there are triples $\{a, b, c\}$ of self-adjoint operators such that

$$c = f(a) = g(b),$$

i.e., c is both a function of a and a function of b. Such a triple is called a Kochen–Specker triple. Crucially, $c = f(a)$ commutes with a and $c = g(b)$ commutes with b, but a need not commute with b. In this way, the functional composition principle does pose conditions on noncommuting operators, too: since $f(v(a)) = v(f(a)) = v(g(b)) = g(v(b))$, the (hypothetical) values $v(a)$ and $v(b)$ assigned to a (respectively, b) are not independent.

The relations induced by Kochen–Specker triples led Isham and Butterfield to introduce the *spectral presheaf*. In their first paper[10] (of four), the self-adjoint operators themselves served as "stages," and in the third paper,[11] the step to commutative subalgebras of $\mathcal{B}(\mathcal{H})$ and the context category was taken. We will focus on the latter. First, consider a single commutative subalgebra V of $\mathcal{B}(\mathcal{H})$. We assume for the moment that V is closed in the norm topology, and hence, V is a C^*-algebra. [For details on the norm topology and C^*-algebras, see, e.g., Ref. 12. In finite dimensions, any subalgebra of $\mathcal{B}(\mathcal{H})$ is a C^*-algebra.] Not surprisingly, there are valuation functions on V_{sa}, the self-adjoint operators in V: every character, that is, every multiplicative linear functional of norm 1,

$$\lambda : V_{\mathrm{sa}} \longrightarrow \mathbb{R}$$
$$a \longmapsto \lambda(a),$$

fulfills both the spectrum rule and the functional composition principle. Conversely, every valuation function is a character of V. Recall from Sec. II A that multiplicative linear functionals of norm 1 on a commutative von Neumann algebra V are exactly the pure states of V. Hence, we consider the set

$$\Sigma(V) := \{\lambda : V_{\mathrm{sa}} \to \mathbb{R} \mid \lambda \text{ is a multiplicative linear functional of norm 1}\}$$

of characters (or pure states) of V, traditionally called the Gelfand spectrum of V.[31] In physical terms, $\Sigma(V)$ is the (pure) state space of the physical system described by the physical quantities in V. As expected, the points of the state space $\Sigma(V)$ correspond exactly with valuation functions on V_{sa}.

If $\tilde{V} \subset V$ is a unital C^*-subalgebra, then every character of $\tilde{V}$ arises as the restriction of some character of V, that is, there is a surjective map

$$\Sigma(V) \longrightarrow \Sigma(\tilde{V}),$$
$$\lambda \longmapsto \lambda|_{\tilde{V}}.$$

Hence, there is a canonical map from the state space of the bigger algebra V to the state space of the smaller algebra $\tilde{V}$. Every valuation function on V can be restricted to a valuation function on $\tilde{V}$.

In infinite dimensions, it is useful to work with commutative von Neumann subalgebras instead of the more general C^*-subalgebras, and we will do so from now on. (In finite dimensions, there is no difference.) Isham, Butterfield, and Hamilton's key idea[10,11] was to combine all the state spaces for commutative subalgebras of a quantum system into one global object. This is the spectral presheaf.

Definition 4. Let $\mathcal{H}$ be a Hilbert space. The spectral presheaf $\underline{\Sigma}$ of the algebra $\mathcal{B}(\mathcal{H})$ is the presheaf over the context category $\mathcal{V}(\mathcal{H}) := \mathcal{V}(\mathcal{B}(\mathcal{H}))$ given

(i)   *on objects: for all commutative von Neumann subalgebras $V \in \mathcal{V}(\mathcal{H})$, let*

$$\underline{\Sigma}_V = \Sigma(V), \quad \text{the Gelfand spectrum of } V;$$

(ii) *on arrows: for all inclusions $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$, let*

$$\underline{\Sigma}(i_{\tilde{V}V}) : \underline{\Sigma}_V \longrightarrow \underline{\Sigma}_{\tilde{V}},$$
$$\lambda \longmapsto \lambda|_{\tilde{V}}.$$

It is clear by construction that the spectral presheaf $\underline{\Sigma}$ is a kind of state space for the quantum system, built from all the state spaces $\Sigma(V)$ of the commuting, compatible parts $V \in \mathcal{V}(\mathcal{H})$ of the noncommutative algebra $\mathcal{B}(\mathcal{H})$ of physical quantities.

As we saw in Sec. II B, for a presheaf, the analog of a point is a global section. What would a global section of the spectral presheaf $\underline{\Sigma}$ be? For every context $V \in \mathcal{V}(\mathcal{H})$, we have to pick one element $\lambda_V \in \underline{\Sigma}_V$, the Gelfand spectrum of V. λ_V is a valuation function for the physical quantities in V, i.e., it assigns a value $\lambda_V(a)$ to all $a \in V_{sa}$ such that the spectrum rule and functional composition hold.

Moreover, if $\dot{V}$ is another commutative subalgebra that contains a, then a is also contained in $\tilde{V} := V \cap \dot{V}$. The value we assign to a in V is $\lambda_V(a)$ and the value we assign to a in $\dot{V}$ is $\lambda_{\dot{V}}(a)$. Moreover, the value we assign to a in $\tilde{V} = V \cap \dot{V}$ is

$$\lambda_{\tilde{V}}(a) = \lambda_V|_{\tilde{V}}(a) = \lambda_V(a)$$

and also

$$\lambda_{\tilde{V}}(a) = \lambda_{\dot{V}}|_{\tilde{V}}(a) = \lambda_{\dot{V}}(a)$$

and hence,

$$\lambda_{\tilde{V}}(a) = \lambda_V(a) = \lambda_{\dot{V}}(a).$$

The structure of a global section, therefore, guarantees that the value assigned to a physical quantity, represented by the self-adjoint operator a, is the same, independent of the context in which it lies. Since also the spectrum rule and the functional composition principle hold, every global section of $\underline{\Sigma}$ would provide a valuation function on all of $\mathcal{B}(\mathcal{H})$. Conversely, a valuation function would give a global section of $\underline{\Sigma}$.

Since the Kochen–Specker theorem shows that there are no valuation functions, i.e., no non-contextual value assignments, Isham, Butterfield, and Hamilton could give the following reformulation:

Theorem 13. *The Kochen–Specker theorem is equivalent to the fact that the spectral presheaf $\underline{\Sigma}(\mathcal{H})$ has no global sections whenever* $\dim(\mathcal{H}) \geq 3$.

In more physical terms, the Kochen–Specker theorem is equivalent to the fact that the quantum state space $\underline{\Sigma}$ has no points. This does not mean, however, that $\underline{\Sigma}$ is "empty," and it still has plenty of subobjects (which are the presheaf analog of subsets). One can just not "focus down" to points, which would be (nonexistent) microstates.

We note that the nonexistence of global sections is not just a consequence of Kochen–Specker but is exactly equivalent. This shows that the context category and the spectral presheaf contain just the right amount of information and that the Kochen–Specker theorem indeed is encoded by our notion of contextuality.

The Kochen–Specker theorem was generalized to von Neumann algebras in Ref. 32.

Theorem 14 *(Generalized Kochen–Specker theorem). Let $\mathcal{N}$ be a von Neumann algebra with no direct summand of type I_1, I_2. Then, there are no valuation functions $v : \mathcal{N}_{\mathrm{sa}} \to \mathbb{R}$.*

The condition that $\mathcal{N}$ has no summand of type I_1, I_2 generalizes the condition that $\dim \mathcal{H} \geq 3$ in the original proof of the theorem.[33,76,77] Since we can easily define the spectral presheaf of a von Neumann algebra $\mathcal{N}$ [simply replace $\mathcal{B}(\mathcal{H})$ by $\mathcal{N}$ and $\mathcal{V}(\mathcal{H})$ by $\mathcal{V}(\mathcal{N})$ in Definition 4], we also have the following reformulation of the generalized Kochen–Specker theorem:

Corollary 15. Let $\mathcal{N}$ be a von Neumann algebra with no direct summand of type I_2, let $\mathcal{V}(\mathcal{N})$ be the context category of $\mathcal{N}$, and let $\underline{\Sigma}$ be its spectral presheaf. The generalized Kochen–Specker theorem is equivalent to the fact that $\underline{\Sigma}$ has no global sections.

V. REFORMULATION OF GLEASON'S THEOREM

A. The Born rule and Gleason's theorem

Gleason's theorem,[2] proven in 1957, shows that the Born rule follows from very modest assumptions. Let $\mathcal{H}$ be a Hilbert space of dimension 3 or greater. Assume that there is a function assigning probabilities to projection operators,

$$\mu : \mathcal{P}(\mathcal{H}) \longrightarrow [0,1],$$

such that

(a) $\mu(1) = 1$,
(b) for all $p, q \in \mathcal{P}(\mathcal{H})$, if $pq = 0$, then $\mu(p + q) = \mu(p) + \mu(q)$.

Condition (a) is the obvious *normalization* condition and (b) is *finite additivity* on mutually orthogonal projections. Clearly, if one aims to have any probabilistic formalism relating to projections (representing propositions about a quantum system), having such a function μ that assigns probabilities to projections is the minimal and natural requirement. There is a built-in non-contextuality condition: every projection p lies in many different contexts, but μ assigns just one probability to p, independently of contexts.

There is an obvious strengthening of finite additivity to infinite families of mutually orthogonal projections, called complete *additivity*:

(b′) for any family $(p_i)_{i \in I}$ of mutually orthogonal projections (i.e., $p_i p_j = \delta_{ij} p_i$ for all $i, j \in I$), it holds that $\mu(\bigvee_{i \in I} p_i) = \sum_{i \in I} \mu(p_i)$.

Note that if the underlying Hilbert space $\mathcal{H}$ is separable, then the index set I is at most countable.[34]

Gleason showed the following theorem, partly answering an earlier problem posed by Mackey.

Theorem 16 *(Gleason's theorem). Let* $\dim \mathcal{H} \geq 3$. *Given a completely additive probability measure μ on projections, there always exists a unique positive trace-class operator with trace* 1, *written ρ_μ, such that*

$$\forall p \in \mathcal{P}(\mathcal{H}) : \mu(p) = \operatorname{tr}[\rho_\mu p].$$

In finite dimensions, ρ_μ is nothing but a density matrix. As mentioned in Sec. II A, this means that every completely additive probability measure on projections determines a unique normal state of $\mathcal{B}(\mathcal{H})$. Conversely, every normal state, equivalently every positive trace-class operator of trace 1 (or, in finite dimensions, every density matrix), ρ determines a unique completely additive probability measure by

$$\mu_\rho : \mathcal{P}(\mathcal{H}) \longrightarrow [0, 1],$$
$$p \longmapsto \operatorname{tr}[\rho p].$$

Obviously, $\mu_{\rho_\mu} = \mu$ and $\rho_{\mu_\rho} = \rho$. Hence, Gleason's theorem justifies the use of density matrices and the Born rule in quantum mechanics. The condition that the Hilbert space is at least three-dimensional is essential, and we will assume $\dim \mathcal{H} \geq 3$ from now on.

In order to understand the power of Gleason's theorem, note that the definition of a completely additive probability measure $\mu : \mathcal{P}(\mathcal{H}) \to [0, 1]$ only poses conditions on mutually orthogonal, hence, commuting projections. In other words, for any family $(p_i)_{i \in I}$ of mutually orthogonal projections, condition (b′) above is a condition within a context V that contains all the p_i, $i \in I$. If there are several contexts that contain all the p_i, it does not matter which context we consider, since probabilities are assigned directly to projections, independently of the contexts they lie in.

For simplicity, let us assume that the Hilbert space $\mathcal{H}$ is finite-dimensional, $\mathcal{H} = \mathbb{C}^n$. Let V be a context of $\mathcal{B}(\mathcal{H}) = M_n(\mathbb{C})$, the complex $n \times n$-matrix algebra. Let $\{p_1, \ldots, p_m\}$ denote the unique set of minimal projections in V. Then, the p_i are mutually orthogonal and V is generated by them, $V = \{p_1, \ldots, p_m\}''$. Every self-adjoint operator $a \in V_{sa}$ is a unique real linear combination of the p_i, that is, $a = \sum_{i=1}^m A_i p_i$. We extend the finitely additive probability measure μ to a function $\mu : V_{sa} \to \mathbb{R}$ by

$$\forall a \in V_{sa} : \mu(a) := \sum_{i=1}^m A_i \mu(p_i).$$

This implies directly that if $r \in \mathbb{R}$, then $\mu(ra) = r\mu(a)$ and if $a, b \in V_{sa}$, then $\mu(a + b) = \mu(a) + \mu(b)$. Hence, $\mu : V_{sa} \to \mathbb{R}$ is a real-linear function. Importantly, if a and b do not commute, it is not obvious at all if $\mu(a + b) = \mu(a) + \mu(b)$ holds or not. A finitely additive probability measure $\mu : \mathcal{P}(\mathcal{H}) \to [0, 1]$ gives a function $\mu : V_{sa} \to \mathbb{R}$ that is *linear in every context V* in a straightforward way, but it is not clear initially why this function should also be linear across contexts, i.e., on noncommuting operators. Traditionally, a function that is linear on commuting operators is called *quasi-linear*.

Gleason's result shows that there always exists a density matrix ρ_μ such that $\mu(p) = \operatorname{tr}[\rho_\mu p]$, and clearly, we also have $\mu(a) = \operatorname{tr}[\rho_\mu a]$ due to linearity of the trace. Crucially, the map

$$\operatorname{tr}[\rho_\mu \_] : \mathcal{B}(\mathcal{H})_{sa} \longrightarrow \mathbb{R}$$
$$a \longmapsto \operatorname{tr}[\rho_\mu a]$$

is linear on *all* (self-adjoint) operators,

$$\forall a, b \in \mathcal{B}(\mathcal{H})_{sa} : \operatorname{tr}[\rho_\mu (a + b)] = \operatorname{tr}[\rho_\mu a] + \operatorname{tr}[\rho_\mu b],$$

which implies that μ is also linear on all operators. Hence, the quasi-linear function μ is, in fact, linear. In this way, Gleason's theorem solves a local-to-global problem, where "local" here means "on commuting operators" (or "within contexts") and global means "on all operators."

By the efforts of many people, Gleason's theorem has been generalized to von Neumann algebras (see Ref. 35 and references therein).

Theorem 17 *(Generalized Gleason's theorem). Let $\mathcal{N}$ be a von Neumann algebra with no direct summand of type I_2, and let $\mu : \mathcal{P}(\mathcal{N}) \to [0, 1]$ be a finitely additive probability measure on the projections of $\mathcal{N}$. There exists a unique state ρ_μ of $\mathcal{N}$ such that*

$$\forall p \in \mathcal{P}(\mathcal{N}) : \mu(p) = \rho_\mu(p).$$

Note that here $\rho_\mu : \mathcal{N} \to \mathbb{C}$ denotes the state itself (i.e., a positive linear functional of norm 1), while before the state was denoted, $\mathrm{tr}[\rho_{\mu-}] : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ and ρ_μ was just the positive trace-class operator (or density matrix). The reason is that the state ρ_μ need not be normal, and hence, there may be no density matrix. In fact, ρ_μ is normal if and only if the probability measure μ is completely additive.

B. The probabilistic presheaf

In order to relate Gleason's theorem (in its generalized form), Theorem 17, more explicitly to contextuality, we consider a certain presheaf that encodes probability assignments to projections. The obvious definition is as follows:

Definition 5. Let $\mathcal{N}$ be a von Neumann algebra with context category $\mathcal{V}(\mathcal{N})$. The probabilistic presheaf $\underline{\Pi}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ is the presheaf given

(i) *on objects: for all $V \in \mathcal{V}(\mathcal{N})$, let*

$$\underline{\Pi}_V := \{\mu_V : \mathcal{P}(V) \to [0,1] \mid \mu_V \text{ is a finitely additive probability measure}\},$$

(ii) *on arrows: for all inclusions $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$, let*

$$\underline{\Pi}(i_{\tilde{V}V}) : \underline{\Pi}_V \longrightarrow \underline{\Pi}_{\tilde{V}}$$
$$\mu_V \longmapsto \mu_V|_{\tilde{V}}.$$

Here, the restriction $\mu_V|_{\tilde{V}}$ of the function $\mu_V : \mathcal{P}(V) \to [0,1]$ to $\mathcal{P}(\tilde{V}) \subset \mathcal{P}(V)$ is simply marginalization.

Note that this is the simplest possible definition of a presheaf built from finitely additive probability measures (FAPMs) on contexts. An element $\mu_V \in \underline{\Pi}_V$ is a FAPM for the projections in V, so it only assigns probabilities to projections in V, not to all projections (unlike the FAPM $\mu : \mathcal{P}(\mathcal{N}) \to [0,1]$ in the generalized Gleason's theorem (Theorem 17), which assigns probabilities to all projections in $\mathcal{N}$).

The probabilistic presheaf $\underline{\Pi}$ can be seen as a generalization of the spectral presheaf $\underline{\Sigma}$ in the following way: at each context $V \in \mathcal{V}(\mathcal{N})$, the component $\underline{\Sigma}_V$ of $\underline{\Sigma}$ is the set of pure states $\lambda : V \to \mathbb{C}$; see Definition 4. In the probabilistic presheaf $\underline{\Pi}$, on the other hand, the component $\underline{\Pi}_V$ is given by finitely additive probability measures $\mu_V : \mathcal{P}(V) \to [0,1]$. The latter are positive linear functionals on V of norm 1, i.e., convex linear combinations of elements in $\underline{\Sigma}_V$. In other words, the elements of $\underline{\Pi}_V$ correspond with mixed states of V, while the elements of $\underline{\Sigma}_V$ correspond with pure states of V, equivalently, extreme points of $\underline{\Pi}_V$.

What about global sections of the probabilistic presheaf? Prima facie, we do not know whether global sections exist or not, but we now show that every quantum state $\rho : \mathcal{N} \to \mathbb{C}$ gives a global section γ_ρ of $\underline{\Pi}$. Define

$$\forall V \in \mathcal{V}(\mathcal{N}) : \ \gamma_\rho(V) := (\rho|_{\mathcal{P}(V)} : \mathcal{P}(V) \to [0,1]).$$

Here, $\rho|_{\mathcal{P}(V)}$ is the restriction of the quantum state to the projections in the context V. Since ρ is linear, $\rho|_{\mathcal{P}(V)}$ is a finitely additive probability measure on the projections in V. If a projection p is contained in a context V and a subcontext $\tilde{V} \subset V$, then

$$\gamma_\rho(\tilde{V})(p) = \rho|_{\mathcal{P}(\tilde{V})}(p) = \rho|_{\mathcal{P}(V)}(p) = \gamma_\rho(V)(p),$$

so γ_ρ is indeed a global section.

Conversely, let γ be a global section of the probabilistic presheaf $\underline{\Pi}$. In every context $V \in \mathcal{V}(\mathcal{N})$, we have a FAPM $\gamma(V) : \mathcal{P}(V) \to [0,1]$ on the projections of V. The restriction maps $\underline{\Pi}(i_{\tilde{V}V})$ guarantee that whenever a context $\tilde{V}$ is contained in another context V, a projection p is assigned the same probability, no matter whether we regard p as a projection in V or in $\tilde{V}$. Hence, a global section γ of the probabilistic presheaf $\underline{\Pi}$ gives a finitely additive probability measure μ on all projections in $\mathcal{N}$. By Gleason's theorem for von Neumann algebras, this determines a unique state ρ_γ of the algebra $\mathcal{N}$, provided $\mathcal{N}$ has no type I_2 summand. The latter condition is akin to the condition that the Hilbert space must be at least three-dimensional.

Before we state Gleason's theorem in its contextual reformulation, we discuss the following slight variation of Definition 5. Note that we may interpret the probability measures $\mu_V \in \underline{\Pi}_V$ as positive operator-valued measures. In fact, by Gelfand duality, every commutative von Neumann algebra $V \in \mathcal{V}(\mathcal{N})$ corresponds with an (extremely disconnected) compact Hausdorff space, whose σ-algebra of open (and closed) sets corresponds with the projection lattice $\mathcal{P}(V)$. Identifying $\mathbb{R}$ with real-valued (1×1)-matrices, μ_V trivially becomes a positive operator-valued measure, and by Naimark's theorem,[36] we can find a dilation of the form $\mu_V = v^* \varphi_V v$, where $v : \mathbb{C} \to \mathcal{K}$ is a bounded linear map into some Hilbert space $\mathcal{K}$ [by scalar multiplication, v corresponds to a vector $v \in \mathcal{K}$ given by $v = v(1)$] and $\varphi_V : \mathcal{P}(V) \to \mathcal{P}(\mathcal{K})$ is an embedding (or spectral measure).[37] In this reading, we obtain a finitely additive probability measure μ_V by setting $\varphi_V|_{\tilde{V}} = \varphi_{\tilde{V}}$ whenever $\tilde{V} \subset V$. It is easy to see that the latter then defines an orthomorphism $\varphi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{K})$, which by Dye's theorem[20,38] lifts to a unique Jordan $*$-homomorphism $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{B}(\mathcal{K})$.

Importantly, collections of dilations over contexts $(\mu_V = v^* \varphi_V v)_{V \in \mathcal{V}(\mathcal{N})}$ still correspond with quantum states $\rho = v^* \Phi v$. In finite dimensions, the latter is easily recognized as a purification of ρ.[39] In particular, restricting to pure states, we may choose $\mathcal{K} = \mathcal{H}$ and $\varphi : \mathcal{P}(\mathcal{H}) \to \mathcal{P}(\mathcal{H})$ the identity map such that $|v\rangle \in \mathcal{H}$ is the pure state corresponding to $\rho(p) = \langle v|p|v \rangle = \mathrm{tr}[|v\rangle\langle v|p]$ for all $p \in \mathcal{P}(\mathcal{N})$. As always, mixed states correspond with convex combinations of pure states; in this sense, applying Naimark's theorem in contexts amounts to a

type of intrinsic convexity condition with respect to the set of pure states (for more details, see Refs. 40 and 41). Taking the latter into account, we refine Definition 5 as follows:

Definition 6. Let $\mathcal{N}$ be a von Neumann algebra with context category $\mathcal{V}(\mathcal{N})$ and $\mathcal{K}$ be a Hilbert space. The dilated probabilistic presheaf $\underline{\Pi}_D : \mathcal{V}(\mathcal{N})^{op} \to \mathbf{Set}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ is the presheaf given

(i) *on objects: for all $V \in \mathcal{V}(\mathcal{N})$, let*
$$(\underline{\Pi}_D)_V := \{\mu_V : \mathcal{P}(V) \to [0,1] \mid \mu_V = v^* \varphi_V v \quad \text{for } v \in \mathcal{K}, \varphi_V : \mathcal{P}(V) \hookrightarrow \mathcal{P}(\mathcal{K}), \text{ and } \mu_V(1) = 1\},$$

(ii) *on arrows: for all inclusions $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$, let*
$$\underline{\Pi}_D(i_{\tilde{V}V}) : (\underline{\Pi}_D)_V \longrightarrow (\underline{\Pi}_D)_{\tilde{V}},$$
$$v^* \varphi_V v = \mu_V \longmapsto \mu_{\tilde{V}} = v^* \varphi_V|_{\tilde{V}} v.$$

One can consider a presheaf that is closely related to the (dilated) probabilistic presheaf $\underline{\Pi}$ ($\underline{\Pi}_D$) but has as component at $V \in \mathcal{V}(\mathcal{N})$ only completely additive probability measures. It is easy to check that if a von Neumann algebra $\mathcal{N}$ has no type I_2 summand, there is a bijective correspondence between normal states on $\mathcal{N}$ and global sections of the normal (dilated) probabilistic presheaf over $\mathcal{V}(\mathcal{N})$. To simplify notation, we will denote the normal (dilated) probabilistic presheaf of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ also by $\underline{\Pi}$ ($\underline{\Pi}_D$).

Theorem 18 *(Generalized Gleason's theorem in contextual form). Let $\mathcal{N}$ be a von Neumann algebra with no direct summand of type I_2. There is a bijective correspondence between quantum states, that is, states on $\mathcal{N}$, and global sections of the (dilated) probabilistic presheaf $\underline{\Pi}$ ($\underline{\Pi}_D$) over $\mathcal{V}(\mathcal{N})$.*

This is our reformulation of Gleason's theorem, which connects it explicitly with contextuality. In contrast to the spectral presheaf, the (dilated) probabilistic presheaf does have global sections and they correspond exactly with quantum states. It is remarkable that the very simple definition of the probabilistic presheaf $\underline{\Pi}$, with FAPMs in every context, connected by the obvious restriction maps in the form of marginalization, suffices to guarantee this (for single systems). In Sec. VI, we will see that FAPMs have to be refined to dilations in contexts, as defined in Definition 6, in order to guarantee a similar correspondence also for composite systems.

As usual, the power of the construction lies in the restriction maps (and, of course, Gleason's theorem). In particular, no further local or global data are needed. In physical terms, there is no need for hidden variables. More importantly, there is no room for hidden variables: as soon as a theory assigns probabilities to all projections in dimension 3 or greater in the obvious way, i.e., finitely additively on orthogonal projections, there exists a quantum state that provides this assignment of probabilities. There are no other (finitely additive) assignments of probabilities to projections apart from those given by quantum states.

Any hidden variables or other extra data could at best give further restrictions. It is worthwhile mentioning the case $\mathcal{N} = \mathcal{B}(\mathcal{H})$ explicitly.

Corollary 19 *(Gleason's theorem in contextual form). Let $\mathcal{B}(\mathcal{H})$ be the algebra of all bounded operators on a Hilbert space $\mathcal{H}$. If $\dim \mathcal{H} \geq 3$, then there is a bijective correspondence between quantum states, that is, states on $\mathcal{B}(\mathcal{H})$, and global sections of the (dilated) probabilistic presheaf $\underline{\Pi}$ ($\underline{\Pi}_D$) over $\mathcal{V}(\mathcal{H})$.*

The fact that Gleason's theorem is closely linked with contextuality in this manner was first observed by Döring,[32] made more explicit by de Groote,[42] and then in a form very similar to the one above by Döring.[43] In a different vein, Gleason's theorem has also been extended to effect algebras,[44,45] which further allows us to cover the two-dimensional case.

VI. BELL'S THEOREM AND CONTEXTUALITY

Bell's seminal paper[4] responds to a long-standing conjecture by Einstein, Podolsky, and Rosen (EPR),[46] who that claim quantum theory is only a statistical version of a more fundamental theory, similar to the relation between thermodynamics and statistical mechanics. Besides the probabilistic nature of quantum theory, this idea is motivated by certain nonlocal features present in the quantum formalism, believed to be resolved within the more fundamental theory. As a response to EPR's famous thought experiment, Bell formalizes EPR's assumption of an underlying space of hidden variables and derives a constraint for the maximal amount of correlations possible in such theories under the additional assumption of locality.[4,47] However, some quantum mechanically predicted and experimentally verified correlations[48-50] do not obey these constraints and, thus, cannot be reproduced by any local hidden variable model.

Pitowsky has pointed out that Bell inequalities are a special case of Boole's "conditions of possible experience"—consistency constraints on correlations under the assumption of a common underlying measure space.[51,52] Here, we extend this perspective in several ways. In Sec. VI A, we identify the assumption of a common underlying measure space with the classical case of trivial physical contextuality and stress the role of composition (of measure spaces) as the relevant notion of locality in Bell inequalities. In Sec. VI B, we use Gelfand duality to show that, in the single-context case, composition of measure spaces and observable algebras coincide. Shifting focus to the latter and their order structure imposed by physical contextuality, in Sec. VI C, we generalize the locality constraint inherent to Bell inequalities from one to many contexts by defining the Bell presheaf as a version of the probabilistic presheaf over product contexts. Surprisingly, global sections of the Bell presheaf correspond with quantum states up to a choice of time orientation in subsystems. This is interpreted as a reformulation and extension of Bell's theorem in contextual form.

A. Correlations in classical theories

1. Classical state spaces

We first give an account of what we mean by a *classical* theory. For our purposes, it will be enough to consider the kinematics and so we start with a set (soon to be upgraded to an algebra) of observables $\mathcal{O}$. We take as a defining property of a classical theory that all its observables are *simultaneously measurable*; from the perspective of physical contextuality, we are thus considering the trivial case of a single context.[53,54,78] Observables $a \in \mathcal{O}$ in classical theories are mathematically represented by functions $f_a : \Sigma \to \mathbb{R}$ on some measure space (Σ, σ, ds). We call Σ the *(single-context) state space* of the theory. Every microstate $s \in \Sigma$ assigns truth values to propositions of the form $'a \in \Delta'$ (read "*the physical quantity a has a value within the Borel subset $\Delta \in \mathbb{R}$*"),

$$\Theta('a \in \Delta', s) := \begin{cases} 1 \text{ if } s \in f_a^{-1}(\Delta), \\ 0 \text{ otherwise.} \end{cases} \tag{1}$$

We can therefore speak of the *value of an observable* $v_s(a)$, *given the state* $s \in \Sigma$ in the intuitive sense, i.e., through evaluation of the corresponding function,

$$v_s(a) := f_a(s). \tag{2}$$

The *valuation functions* $v_s : \mathcal{O} \to \mathbb{R}$ in Eq. (2), which we already discussed in connection with the Kochen–Specker theorem in Sec. IV, are defined on all observables; in other words, every observable has an intrinsic (sharp) value in every state. [Note that the spectrum rule, $v_s(a) \in \mathrm{sp}(a) = \mathrm{Im}(f_a)$, is trivially satisfied.] The observation that all observables simultaneously take deterministic values justifies to model *physical states* by points in some space Σ and observables by functions $f_a : \mathcal{O} \to \mathbb{R}$ in the first place. Of course, this inductive reasoning has to be revisited for non-classical theories, that is, theories with non-trivial physical contextuality. We will do so in Secs. VI B and VI C.

It is natural to equip the set of observables $\mathcal{O}$ with the structure of an algebra.[55] It is straightforward to extend the definition of valuation functions in Eq. (2) to this extra algebraic structure: for all $a, b \in \mathcal{O}, r \in \mathbb{R}$ and $s \in \Sigma$, we set

$$v_s(a \cdot b) := f_a(s) \cdot f_b(s), \quad v_s(a + b) := f_a(s) + f_b(s), \quad v_s(ra) := r f_a(s). \tag{3}$$

From this perspective, classical states $s \in \Sigma$ correspond with *algebra homomorphisms* $v_s : \mathcal{O} \to \mathbb{R}$. Note that when $\mathcal{O} = \mathcal{N}$ is a commutative von Neumann algebra, Eq. (3) holds as a consequence of the functional composition principle $v_s(f(a)) = f(v_s(a))$, in which $f : \mathbb{R} \to \mathbb{R}$ is a continuous function and $a, f(a) \in \mathcal{N}_{\mathrm{sa}}$ are self-adjoint operators (see Sec. IV).

Given two subsystems with measure spaces $(\Sigma_1, \sigma_1, ds_1)$ and $(\Sigma_2, \sigma_2, ds_2)$, the composite state space is defined as the Cartesian product $\Sigma_{1\,\&\,2} := \Sigma_1 \times \Sigma_2$ with product σ-algebra $\sigma_{1\,\&\,2}$ generated by elements $B_1 \times B_2, B_1 \in \sigma_1, B_2 \in \sigma_2$ and product measure $ds_{1\,\&\,2} := ds_1 \times ds_2$ satisfying the condition[56]

$$(ds_1 \times ds_2)(B_1 \times B_2) = ds_1(B_1) \cdot ds_2(B_2). \tag{}$$

In a similar way, we obtain composite state spaces with multiple subsystems. If the algebra $\mathcal{O}$ of the composite system is generated from the algebras of its subsystems, i.e., by functions $f_a : \Sigma \to \mathbb{R}^n$ on the composite state space $\Sigma = \times_{i=1}^n \Sigma_i$ (see Sec. VI B), then evaluation on elements $s \in \Sigma$ again yields algebra homomorphisms similarly to Eq. (3). Hence, we obtain composite valuation functions $v_s : \mathcal{O} \to \mathbb{R}^n$ from the obvious generalization of Eq. (2) to composite observables. Moreover, we obtain a generalization of the truth values in Eq. (1) by considering tuples $\mathbf{a} = (a_1, \ldots, a_n) \in \mathcal{O}$ with $a_i \in \mathcal{O}_i$ for $i \in \{1, \ldots, n\}$ as well as functions $f_{\mathbf{a}} : \Sigma \to \mathbb{R}^n, f_{\mathbf{a}}(s) := (f_{a_1}(s_1), \ldots, f_{a_n}(s_n))$ with $s \in \Sigma = \times_{i=1}^n \Sigma_i$. The truth value of the proposition $'\mathbf{a} \in \Delta'$ with the Borel set $\Delta := \times_{i=1}^n \Delta_i$ is defined as

$$\Theta('\mathbf{a} \in \Delta', s) := \begin{cases} 1 \text{ if } s \in f_{\mathbf{a}}^{-1}(\Delta) \\ 0 \text{ otherwise} \end{cases} = \begin{cases} 1 \text{ if } s_i \in f_{a_i}^{-1}(\Delta_i) \; \forall i \\ 0 \text{ otherwise} \end{cases} = \prod_{i=1}^n \Theta('a_i \in \Delta_i', s_i). \tag{4}$$

2. Factorizability of joint probability distributions

The spectrum rule [Eq. (2)] and the algebraic relations in Eq. (3) are characteristic of pure states. Mixed states are convex mixtures of pure states, i.e., probability measures $\mu : \Sigma \to \mathbb{R}$,

$$\int_\Sigma ds\, \mu(s) = 1, \quad \mu(s) \geq 0 \quad \forall s \in \Sigma.$$

The probability for the event corresponding to the Borel set $\Delta \subset \mathbb{R}$ when measuring the observable $a \in \mathcal{O}$ of a system in the mixed state μ is given by[57]

$$P(\Delta \mid a) = \int_{\{s \in \Sigma \mid v_s(a) \in \Delta\}} ds\, \mu(s) = \int_{f_a^{-1}(\Delta)} ds\, \mu(s) = \int_\Sigma ds\, \mu(s)\, \Theta('a \in \Delta', s). \tag{5}$$

Note that in the last step we have used the indicator function $\Theta('a \in \Delta', s)$ in Eq. (1). For instance, the probability for obtaining a particular outcome A corresponds to $\Delta_A := \{A\}$. Analogously, for joint probability distributions on a bipartite system (represented by the product of measure spaces $\Sigma = \Sigma_1 \times \Sigma_2$), we obtain

$$P(A, B \mid a, b) = \int_{\Sigma} ds \, \mu(s) \, \Theta('a \in \Delta_A', s_1) \cdot \Theta('b \in \Delta_B', s_2). \tag{6}$$

This condition on joint probability distributions is called *factorizability*. From it, one derives the Bell inequalities in the usual way.[4,58]

Note that Eq. (6) expresses merely the splitting of supports of indicator functions $\Theta('a \in \Delta', s)$. We, thus, conclude that factorizability is a consequence of the following two assumptions:

(a) trivial physical contextuality, i.e., a single common underlying measure space, and
(b) the composite system is given by the Cartesian product of individual measure spaces.

Assumption (b) is the natural way to compose independent (single-context) measure spaces. It encodes the locality condition in the derivation of Bell inequalities, which we identify with local consistency conditions on correlations, interpreted over a common single composite measure space.[51,59] In turn, the violation of Bell inequalities, thus, becomes a no-go-result for the existence of (a joint probability distribution in) a composite classical state space.

We remark that the above argument is not restricted to outcome deterministic models, and the same conclusion obtains if not all s are (experimentally) distinguishable.[60] Note also that we have not yet specified the algebra of functions. For instance, the above results apply to commutative von Neumann as well as C^*-algebras (cf. Ref. 61). In Sec. VI B, we extend the constraints on correlations in Eq. (6) by relaxing assumption (a), i.e., by allowing for non-trivial physical contextuality.

B. Locality and composition of contexts

In Sec. VI A, we have argued that factorizability can be seen as a consequence of the assumption of a common underlying measure space, which by locality decomposes as the product of individual measure spaces. However, by the Kochen–Specker theorem (Theorem 12), this picture is hardly available for quantum systems.[62,79] In order to impose local "conditions of possible experience" (in the sense of Refs. 51 and 52) also for quantum systems, in this section, we shift perspective from measure spaces to observable algebras and their partial order of contexts. In particular, we define a notion of composition for the latter, which extends the locality constraint in Bell inequalities beyond the classical case, i.e., to systems with non-trivial physical contextuality.

1. Composition of classical state spaces

Recall that in Sec. VI A we defined composition of classical systems in terms of their state spaces, namely, via the product of the corresponding measure spaces. Alternatively, we can define the composite state space in terms of the corresponding algebras of observables.

To this end, we represent observables as measurable functions in a commutative von Neumann algebra $\mathcal{N}$. By Gelfand duality, we may identify $\mathcal{N}$ with a measure space over its Gelfand spectrum $\Sigma(\mathcal{N})$.[63,80] Recall from Theorem 18 that global sections of the probabilistic presheaf $\underline{\Pi}$ correspond with (mixed) states on $\mathcal{N}$, i.e., $\mathcal{S}(\mathcal{N}) \cong \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N})))$. Furthermore, if we define the product of two commutative von Neumann algebras $\mathcal{N}_1, \mathcal{N}_2$ in terms of the (spatial) tensor product $\mathcal{N}_1 \otimes \mathcal{N}_2$,[64] it follows that

$$\Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2))) = \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_1 \otimes \mathcal{N}_2))) \cong \mathcal{S}(\mathcal{N}_1 \otimes \mathcal{N}_2), \tag{7}$$

where $\underline{\Pi}$ denotes the normal probabilistic presheaf. It is easy to see that Eq. (7) is not true for noncommutative von Neumann algebras (see also comment below Definition 7). Nevertheless, it holds in every commutative von Neumann subalgebra $V \subset \mathcal{N}$, i.e., in every context $V \in \mathcal{V}(\mathcal{N})$. This suggests to extend the locality assumption in Bell inequalities by evaluating it with respect to *all product contexts*.

2. Composite context category

Let $\mathcal{N}_1, \mathcal{N}_2$ be generally noncommutative von Neumann algebras and $\mathcal{V}(\mathcal{N}_1), \mathcal{V}(\mathcal{N}_2)$ be the respective context categories. We define the composite context category by

$$\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2} := (\mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2), \subset_{1 \& 2}), \tag{8}$$

where for all $\tilde{V}_1, V_1 \in \mathcal{V}(\mathcal{N}_1), \tilde{V}_2, V_2 \in \mathcal{V}(\mathcal{N}_2): (\tilde{V}_1, \tilde{V}_2) \subset_{1 \& 2} (V_1, V_2) :\Leftrightarrow (\tilde{V}_1 \subset_1 V_1, \tilde{V}_2 \subset_2 V_2)$. If $\mathcal{N}_1 = \mathcal{B}(\mathcal{H}_1), \mathcal{N}_2 = \mathcal{B}(\mathcal{H}_2)$, we also write $\mathcal{V}_{\mathcal{H}_1 \& \mathcal{H}_2}$ for $\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2}$. Note that, by Eq. (7), we may identify every composite context $(V_1, V_2) := V_1 \otimes V_2$ with the (spatial) tensor product of the respective commutative von Neumann subalgebras. By comparison with the derivation in Sec. VI A, Eq. (8) is a generalization of the locality constraint in Bell inequalities from one to many contexts.

With $\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2}$ as the base category, we may define the normal probabilistic presheaf $\underline{\Pi}(\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2})$. Each component $\underline{\Pi}_{(V_1, V_2)}$ consists of completely additive probability measures $\mu : \mathcal{P}(V_1 \otimes V_2) \to [0, 1]$ together with the obvious restriction maps. More importantly, we define the *Bell presheaf* as the normal dilated probabilistic presheaf $\underline{\Pi}_D(\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2})$ according to Definition 6. This subtle difference plays a crucial role

in Lemma 20 below (cf. Refs. 39–41). Physically, it represents the idea that we can think of a system as part of a larger system (e.g., including an environment).

Definition 7. Let $\mathcal{N}_1, \mathcal{N}_2$ be a von Neumann algebras with context category $\mathcal{V}(\mathcal{N}_1)$, $\mathcal{V}(\mathcal{N}_2)$, respectively. Then, we call the normal dilated probabilistic presheaf $\underline{\Pi}_D(\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2})$ over the product context category $\mathcal{V}_{1 \& 2} := \mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2)$ the Bell presheaf of $\mathcal{N}_1$ and $\mathcal{N}_2$.

Definition 7 provides a natural setting to study extensions of Bell inequalities, i.e., consistency constraints on correlations in different measure spaces—related to one another by physical contextuality. Importantly, we point out that $\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2}$ assumes much less structure than the (spatial) tensor product of (noncommutative) von Neumann algebras $\mathcal{N}_1 \otimes \mathcal{N}_2$. In particular, $\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2}$ contains far fewer contexts than $\mathcal{V}(\mathcal{N}_1 \otimes \mathcal{N}_2)$, namely, only product contexts.

C. The Bell presheaf

For the sake of clarity of argument, we will restrict the present discussion to the case $\mathcal{N}_i = \mathcal{B}(\mathcal{H}_i)$, $i = 1, 2$, with $\dim(\mathcal{H}_i) \geq 3$ finite, and refer to Ref. 39 for the general case. Recall from Gleason's theorem in contextual form (Theorem 18) that quantum states of the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ bijectively correspond with global sections of the probabilistic presheaf $\underline{\Pi}(\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2))$. Since $\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is a much richer poset than $\mathcal{V}_{\mathcal{H}_1 \& \mathcal{H}_2}$, there are many more restriction maps in the (dilated) probabilistic presheaf $\underline{\Pi}_D(\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2))$ than in the Bell presheaf $\underline{\Pi}(\mathcal{V}_{\mathcal{H}_1 \& \mathcal{H}_2})$. It is easy to see that every global section of $\underline{\Pi}(\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2))$, that is, every quantum state, induces a global section of $\underline{\Pi}(\mathcal{V}_{\mathcal{H}_1 \& \mathcal{H}_2})$, but it is not clear *a priori* whether the converse holds. The Bell presheaf $\underline{\Pi}(\mathcal{V}_{\mathcal{H}_1 \& \mathcal{H}_2})$ could potentially have many more global sections than those corresponding with quantum states.[65,81,82] Remarkably, this is not the case. In order to see this, the following lemma is crucial; for details, we refer to Refs. 39 and 41.

Lemma 20. Let $\mathcal{H}_i$ be Hilbert spaces with $\dim(\mathcal{H}_i) \geq 3$ finite, $\mathcal{B}(\mathcal{H}_i)$ be the algebras of physical quantities, and $\mathcal{V}(\mathcal{H}_i)$ be the corresponding context categories. Then, for every global section $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{\mathcal{H}_1 \& \mathcal{H}_2}))$ of the Bell presheaf in Definition 7, there exists a unique linear map $\phi^\gamma : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$. Moreover, there exists a Hilbert space $\mathcal{K}$, a linear map $v : \mathcal{H}_2 \to \mathcal{K}$, and a Jordan $*$-homomorphism $\Phi^\gamma : \mathcal{J}(\mathcal{H}_1) \to \mathcal{J}(\mathcal{K})$ such that

$$\phi^\gamma = v^* \Phi^\gamma v.$$

ϕ^γ in Lemma 20 is a decomposable map.[66] In contrast, by Choi's theorem,[67] every density matrix on the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ corresponds with a completely positive, trace-preserving map $\phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, and by Stinespring's theorem,[68] every such completely positive map ϕ is of the form $\phi = v^* \Phi v$ with $v : \mathcal{H}_2 \to \mathcal{K}$ a linear map and $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{K})$ a $*$-homomorphism. It follows that a global section of the Bell presheaf $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{\mathcal{N}_1 \& \mathcal{N}_2}))$ corresponds with a quantum state if and only if the Jordan $*$-homomorphism in Lemma 20 lifts to a $*$-homomorphism.

Not every Jordan $*$-homomorphism is also $*$-homomorphism. It turns out that for the special case of $\mathcal{N} = \mathcal{B}(\mathcal{H})$, there are exactly two ways to lift a Jordan algebra $\mathcal{J}(\mathcal{H}) = (\mathcal{B}(\mathcal{H}), \circ)$ to a von Neumann algebra: by augmenting the symmetric product (anticommutator) to an associative product $a \cdot_\pm b = \frac{1}{2}\{a, b\} \pm \frac{1}{2}[a, b]$. Moreover, for every Jordan $*$-homomorphism $\Phi : \mathcal{J}(\mathcal{H}_1) \to \mathcal{J}(\mathcal{K})$, it holds $\Phi([a, b]) = \pm[\Phi(a), \Phi(b)]$, whereas Φ is a $*$-homomorphism only if it also preserves the commutator, $\Phi([a, b]) = [\Phi(a), \Phi(b)]$ for all $a, b \in \mathcal{B}(\mathcal{H}_1)$ (cf. Ref. 69). The difference between the two associative algebras with products $\cdot_\pm$ has a clear physical interpretation in terms of the direction of time. This can be made precise in the form of time orientations on the context category. (For more details, we refer to Refs. 41 and 70 and specifically.[21,22])

Definition 8. Let $\mathcal{H}$ be a Hilbert space, $\mathcal{B}(\mathcal{H})$ be the algebra of physical quantities, and $\mathcal{V}(\mathcal{H})$ be the corresponding context category. The canonical time orientation on the context category is the map into the automorphisms on $\mathcal{V}(\mathcal{H})$, denoted by $\mathrm{Aut}(\mathcal{V}(\mathcal{H}))$,

$$\begin{aligned} \Psi : \mathbb{R} \times \mathcal{J}(\mathcal{H})_{\mathrm{sa}} &\longrightarrow \mathrm{Aut}(\mathcal{V}(\mathcal{H})), \\ (t, a) &\longmapsto e^{ita} V e^{-ita}. \end{aligned} \quad (9)$$

The context category $\mathcal{V}(\mathcal{H})$ together with the canonical time orientation Ψ on it is called the time-oriented context category and denoted $\widetilde{\mathcal{V}(\mathcal{H})} = (\mathcal{V}(\mathcal{H}), \Psi) \cong \mathcal{B}(\mathcal{H})$.

By Theorem 8, every order automorphism on $\mathcal{V}(\mathcal{H})$ corresponds to conjugation by a unitary or anti-unitary operator. Recall that every anti-unitary can be decomposed into the time-reversal operator and a unitary operator, where the time-reversal operator causes a sign change of the time parameter $t \in \mathbb{R}$ in the canonical evolution in Eq. (9).

Moreover, by differentiation of Eq. (9), time-reversal corresponds with a change in sign for the commutator in the two associative algebras with products $\cdot_\pm$, whose common Jordan algebra is $\mathcal{J}(\mathcal{H})$. In line with Definition 8, we denote by Ψ^* the reverse time orientation on $\mathcal{V}(\mathcal{H})$, where Ψ^* is given by the inversion $t \mapsto -t$ in Eq. (9), i.e., $\Psi^*(t, a) = \Psi(-t, a)$.[71,83] Succinctly, time orientations encode the forward time direction in a quantum system.

We remark that for general Jordan algebras there are more ways to lift them to von Neumann algebras and, thus, also more possible time orientations on $\mathcal{N}$.[72]

Next, we introduce a notion of *time-oriented global sections* with respect to the relative time orientation $\Psi = (\Psi_1^*, \Psi_2)$ between systems $(\mathcal{V}(\mathcal{H}_1), \Psi_1^*)$ and $(\mathcal{V}(\mathcal{H}_2), \Psi_2)$. We remark that this relative orientation is implicit in Choi's theorem[67] (for details, see Ref. [70]).

Definition 9. Let $\mathcal{H}_i$, $i = 1, 2$, be Hilbert spaces, $\mathcal{B}(\mathcal{H}_i)$ be the algebras of physical quantities on either subsystem, and $\widetilde{\mathcal{V}(\mathcal{H}_i)}$ be the corresponding context categories with respective canonical time orientations Ψ_i. A global section of the Bell presheaf $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{1\&2}))$ is called time-oriented with respect to $\Psi = (\Psi_1^*, \Psi_2)$ if

$$\forall t \in \mathbb{R}, a \in \mathcal{J}(\mathcal{H}_1)_{\mathrm{sa}} : \quad \Phi^\gamma \circ \Psi_1^*(t, a) = \Psi_2'(t, \Phi^\gamma(a)) \circ \Phi^\gamma,$$

where Φ^γ is the Jordan *-homomorphism in Lemma 20 and Ψ_2' is the unique time-orientation on $\mathcal{B}(\mathcal{K})$ induced by the canonical time orientation Ψ_2 on $\mathcal{B}(\mathcal{H}_2)$. We denote by $\Gamma(\underline{\Pi}(\widetilde{\mathcal{V}(\mathcal{H}_1)}^* \times \widetilde{\mathcal{V}(\mathcal{H}_2)}))$ the set of time-oriented global sections with respect to $\Psi = (\Psi_1^*, \Psi_2)$.

Combining the above, we arrive at a contextual reformulation and extension of Bell's theorem. For the proof and more details, see Refs. [39] and [41].

Theorem 21 (Bell's theorem in contextual form). Let $\mathcal{H}_i$, $i = 1, 2$, be Hilbert spaces $\dim(\mathcal{H}_i) \geq 3$ finite, $\mathcal{B}(\mathcal{H}_i)$ be the algebra of physical quantities, $\widetilde{\mathcal{V}(\mathcal{H}_i)}$ be the corresponding context categories with respective canonical time orientations Ψ_i, and $\mathcal{S}(\mathcal{H}_i) \cong \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{B}(\mathcal{H}_i))))$ be the corresponding state spaces. Then, the state space of the composite system is given by

$$\mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2) \cong \Gamma(\underline{\Pi}(\widetilde{\mathcal{V}(\mathcal{H}_1)}^* \times \widetilde{\mathcal{V}(\mathcal{H}_2)})).$$

Moreover, for commutative von Neumann algebras $\mathcal{N}_1, \mathcal{N}_2$,

$$\mathcal{S}(\mathcal{N}_1 \otimes \mathcal{N}_2) \cong \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2))).$$

Theorem 21 evaluates the constraints on correlations bound by physical contextuality and locality in the form of context composition. It bounds classical correlations by Bell inequalities [cf. Eq. (6) in the commutative (single-context) case as a consequence of Eq. (7)]. More generally, Theorem 21 classifies correlations bound by the natural generalization of Bell inequalities from one to many contexts and their order relations given by inclusion (coarse-graining). It is in this sense that Theorem 21 is a reformulation and extension of Bell's theorem in contextual form. Remarkably, the resulting correlations already correspond with quantum states up to the relative time orientation between local subsystems in Definition 9. The latter is a genuine quantum phenomenon, which is intimately related to entanglement.[73]

The proof of Theorem 21 as presented here is given in Refs. [40] and [41], together with a thorough discussion on the relation between contextuality and no-signaling. The generalization to general von Neumann algebras [replace $\mathcal{B}(\mathcal{H}_i)$ with $\mathcal{N}_i$ in Theorem 21] is proved in Ref. [39], where it is discussed in reference to Gleason's theorem and as a generalization thereof to composite systems. We remark that Theorem 21 provides a (first step towards a) notion of composition in the topos approach to quantum theory.

Finally, we note that we have defined locality in the form of composition of local observable algebras in Eq. (8). This is somewhat reminiscent of the construction of the local observable algebra in algebraic quantum field theory[74] and motivates to further generalize Theorem 21 to multiple systems. We leave this and similar problems for future research.

VII. CONCLUSION AND OUTLOOK

In this article, we have shown that important structural theorems of quantum theory—Wigner's theorem, Gleason's theorem, the Kochen–Specker theorem, and Bell's theorem—fundamentally relate to contextuality.

Wigner's theorem can be rephrased in terms of automorphisms on the partially ordered set of contexts, that is, maps preserving the context order, which are implemented by conjugation with unitary or anti-unitary operators. Hence, instead of demanding transition probabilities between pure states to be preserved, one can equivalently demand the order on contexts to be preserved instead.

The Kochen–Specker theorem is equivalent to the fact that the spectral presheaf has no global sections: given a "local" pure state in every context (each such state assigns sharp values to all observables in its context), there is no way of fitting these together in a consistent way. In other words, there are no dispersion-free quantum states. The "fitting together" here refers to the non-contextuality condition asserting that if an observable is contained in different contexts, then the value assigned to it by the different pure states must be the same.

Gleason's theorem answers a similar local-to-global problem, yet instead of valuation functions, it considers measures on the physical quantities in a quantum system. From the perspective of presheaves, this is easily achieved by replacing pure states with mixed states *locally*, that is, in every context, and by extending the restriction maps (from the spectral preheaf) to all probability measures, which, thus, become marginalization constraints. Again, one asks for global sections, that is, probability assignments that are consistent *globally* or across contexts. In contrast to the spectral presheaf, global sections do exist in the case of the probabilistic presheaf and by Gleason's theorem bijectively correspond with quantum states (density matrices in finite dimensions). Gleason's theorem, therefore, lifts quasi-linearity of probability measures in contexts to linearity on states.

Finally, Bell's theorem attains a reformulation and substantial extension in terms of (non-trivial) contextuality. Bell inequalities naturally arise as constraints on correlations over a common underlying measure space, where locality demands that it is of composite form, i.e., a product of individual measure spaces. From this perspective, Bell inequalities are restricted to individual contexts only. Our contextual

reformulation of Bell's theorem extends Bell inequalities to many contexts. Locality enters as a notion of composition for contexts, and correlations satisfying the collection of all Bell inequalities over the composite context category become global sections of the Bell presheaf. Surprisingly, the latter correspond with quantum states up to a choice of time orientation in subsystems.

ACKNOWLEDGMENTS

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

[1] E. P. Wigner, *Gruppentheorie und Ihre Anwendung auf die Quantenmechanik der Atomspektren* (Friedrich Vieweg und Sohn, Braunschweig, 1931), pp. 251–254.

[2] A. M. Gleason, J. Math. Mech. **6**, 885 (1957).

[3] We use "physical quantity" synonymously with "observable."

[4] J. S. Bell, Phys. Phys. Fiz. **1**, 195 (1964).

[5] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).

[6] N. Bohr, in *Albert Einstein: Philosopher-Scientist*, Library of Living Philosophers Vol. VII, edited by P. A. Schilpp (Open Court, 1998), pp. 199–241.

[7] See also Ref. 75 about Grete Hermann's lucid early contribution in interpreting Bohr's notion of complementarity and her influence on her contemporaries and the subject in general.

[8] R. B. Griffiths, Philos. Trans. R. Soc., A **377**, 20190033 (2019).

[9] J. B. Conway and S. Kochen, Found. Phys. **36**, 1441 (2006).

[10] C. J. Isham and J. Butterfield, Int. J. Theor. Phys. **37**, 2669 (1998).

[11] J. Hamilton, C. J. Isham, and J. Butterfield, Int. J. Theor. Phys. **39**, 1413 (2000).

[12] R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras. Volume I, Elementary Theory* (Academic Press, 1983).

[13] R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras. Volume II, Advanced Theory* (Academic Press, 1983).

[14] M. Takesaki, *Theory of Operator Algebras I* (Springer-Verlag, New York, 1979).

[15] A. Döring and C. J. Isham, in *New Structures for Physics*, Lecture Notes in Physics Vol. 813, edited by B. Coecke (Springer, 2011), pp. 753–937.

[16] S. Mac Lane, *Categories for the Working Mathematician*, Graduate Texts in Mathematics Vol. 5 (Springer-Verlag, 1971).

[17] J. Harding and M. Navara, Order **28**, 549 (2011).

[18] C. Heunen, N. P. Landsman, and B. Spitters, Commun. Math. Phys. **291**, 63 (2009).

[19] G. Cassinelli, E. de Vito, P. Lahti, and A. Levrero, *The Theory of Symmetry Actions in Quantum Mechanics* (Springer, 2004).

[20] H. A. Dye, Ann. Math. **61**, 73 (1955).

[21] A. Döring, "Two new complete invariants of von Neumann algebras," arXiv:1411.5558 [math.OA] (2014).

[22] E. M. Alfsen and F. W. Shultz, *State Spaces of Operator Algebras: Basic Theory, Orientations, and C^*-Products* (Birkhäuser, Boston, 2001).

[23] K. Landsman and B. Lindenhovius, in *Reality and Measurement in Algebraic Quantum Theory*, edited by M. Ozawa, J. Butterfield, H. Halvorson, M. Rédei, Y. Kitajima, and F. Buscemi (Springer, Singapore, 2018), pp. 97–118.

[24] A. Döring and J. Harding, Houston J. Math. **42**, 559 (2016).

[25] Strictly speaking, because $\underline{1}$ is a contravariant functor, we encode the opposite order, hence the poset $\mathcal{V}(\mathcal{N})^{\mathrm{op}}$, but this gives exactly the same amount of information as $\mathcal{V}(\mathcal{N})$.

[26] It is less clear which positive statement can be given and whether "contextual value assignments" exist or would even make sense.

[27] C. J. Isham and J. Butterfield, Int. J. Theor. Phys. **38**, 827 (1999).

[28] C. J. Isham and J. Butterfield, Int. J. Theor. Phys. **41**, 613 (2002).

[29] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[30] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932); *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955).

[31] The Gelfand spectrum $\Sigma(V)$ is equipped with the *Gelfand topology*, which is the topology of pointwise convergence.

[32] A. Döring, Int. J. Theor. Phys. **44**, 139 (2005).

[33] The type I_1 factors in the factor decomposition of $\mathcal{N}$ correspond with its commutative part. In general, commutative von Neumann algebras admit many valuation functions.[32,76,77]

[34]There is an obvious notion of countable additivity: *complete additivity* reduces to countable additivity for a separable and to finite additivity for a finite-dimensional Hilbert space.

[35]S. Maeda, Rev. Math. Phys. **1**, 235 (1989).

[36]M. A. Naimark, C. R. Dokl. Acad. Sci. URSS **41**, 359 (1943).

[37]Note that we dropped the dependence on contexts for the Hilbert space $\mathcal{K}$ and the linear map $v : \mathbb{C} \to \mathcal{K}$. This can be done by choosing $\mathcal{K}$ sufficiently large,[36,68] and by absorbing the context dependence on v into φ_V. Alternatively, one can impose constraints under restriction as in Definition 6 for both v_V and φ_V individually.

[38]L. J. Bunce and J. D. M. Wright, Expositiones Math. **11**, 91 (1993).

[39]M. Frembs and A. Döring, arXiv:2205.00493 [quant-ph] (2022).

[40]M. Frembs and A. Döring, arXiv:1910.09596 [quant-ph] (2019).

[41]M. Frembs and A. Döring, arXiv:2204.11471 [quant-ph] (2022).

[42]H. F. de Groote, "Observables IV: The presheaf perspective," arXiv:0708.0677 [math-ph] (2007).

[43]A. Döring, "Flows on generalised Gelfand spectra of nonabelian unital C^*-algebras and time evolution of quantum systems," arXiv:1212.4882 [math.OA] (2012).

[44]P. Busch, Phys. Rev. Lett. **91**, 120403 (2003).

[45]B. Jacobs and J. Mandemaker, Electron. Proc. Theor. Comput. Sci. **95**, 143–182 (2012).

[46]A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[47]J. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge University Press, 2004).

[48]A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981).

[49]M. Giustina *et al.*, Phys. Rev. Lett. **115**, 250401 (2015).

[50]L. K. Shalm *et al.*, Phys. Rev. Lett. **115**, 250402 (2015).

[51]I. Pitowsky, "From George Boole to John Bell—The origins of Bell's inequality," in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, edited by M. Kafatos (Springer, Dordrecht, The Netherlands, 1989), pp. 37–49.

[52]I. Pitowsky, Br. J. Philos. Sci. **45**, 95 (1994).

[53]D. A. Edwards, Synthese **42**, 1 (1979).

[54]We are not considering additional non-commuting operators, such as in the Koopman–von Neumann formalism for classical mechanics (see Ref. 78, for instance).

[55]We include the "trivial" observable $e \in \mathcal{O}$ represented by the constant function $f_e = 1$. This observable simply asks the question "Is the system there?" and the answer is always "yes." When $\mathcal{O} = \mathcal{N}$ is a commutative von Neumann algebra, e is the identity.

[56]A product measure always exists. It is unique if the individual measures are σ-finite, which is the case for commutative von Neumann algebras on separable Hilbert spaces.

[57]Here, it is natural to assume f to be a measurable function. In Secs. VI B and VI C, we will do so, by taking $\mathcal{O}$ to be a von Neumann algebra.

[58]J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[59]A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[60]In particular, it also holds if no dispersion-free states (truth values) exist, e.g., note that commutative von Neumann algebras $\mathcal{N}_1, \mathcal{N}_2$ admit dispersion-free states if and only if they possess a mimimal projection.[76,77]

[61]J. Baez, Lett. Math. Phys. **13**, 135 (1987).

[62]For a related discussion on the connection between contextuality and Bell's theorem, see, e.g., Ref. 79. We remind the reader that in this treatment we exclude the two-dimensional case.

[63]For technical details on the Gelfand duality for commutative von Neumann algebras, see Ref. 80.

[64]In this case, every normal product state extends to a unique normal state on the composite system, and the normal state space of the composite system is generated from normal product states (Propositions 11.27 and 11.28 in Ref. 13).

[65]In contrast, if in Definition 7 we replace the dilated probabilistic presheaf with the probabilistic presheaf, the resulting Bell presheaf contains many global sections that do not correspond with quantum states.[81,82] For more details, see Refs. 40 and 41.

[66]E. Stormer, Proc. Am. Math. Soc. **86**, 402 (1982).

[67]M.-D. Choi, Linear Algebra Appl. **10**, 285 (1975).

[68]W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).

[69]R. V. Kadison, Ann. Math. **54**, 325 (1951).

[70]M. Frembs, "Variations on the Choi-Jamiołkowski isomorphism" (unpublished) (2022).

[71]The canonical time orientation $\Psi = \exp \circ \psi$ in Definition 8 is the exponential of the (canonical) dynamical correspondence on $\mathcal{B}(\mathcal{H})$ (cf. Ref. 83). The reverse time orientation $\Psi^* = * \circ \exp \circ \psi$.

[72]The notion of time-oriented presheaves over the context category was introduced in Ref. 21.

[73]M. Frembs, "Entanglement and the arrow of time" (unpublished) (2022).

[74]R. Haag and D. Kastler, J. Math. Phys. **5**, 848 (1964).

[75]E. M. Crull and G. Bacciagaluppi, *Grete Hermann–Between Physics and Philosophy* (Springer, Berlin, 2016).

[76]S. P. Gudder, Proc. Am. Math. Soc. **19**, 319 (1968).

[77]R. J. Plymen, Il Nuovo Cimento A **54**, 862 (1968).

[78]P. Morgan, Ann. Phys. **414**, 168090 (2020).

[79]N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).

[80]D. Pavlov, J. Pure Appl. Algebra **226**, 106884 (2022).

[81]M. Kläy, C. Randall, and D. Foulis, Int. J. Theor. Phys. **26**, 199 (1987).

[82]N. R. Wallach, "An unentangled Gleason's theorem," in *Quantum Information and Computation* (American Mathematical Society, 2002), Vol. 305, pp. 291–298.

[83]E. M. Alfsen and F. W. Shultz, Proc. Natl. Acad. Sci. U. S. A. **95**, 6596 (1998).