STUDIES IN STRONG CONTEXTUALITY

AJ DEROSA

Abstract. We survey the work due to Isham, Döring, Abramsky, and others that uses presheaves and related methods to understand strong contextuality as it arises in quantum mechanics and elsewhere. In particular, we make explicit the connection between the Kochen-Specker theorem in topos quantum theory and the existence of certain strongly contextual empirical models, which in turn relate to cohomological obstructions. We explore as well how the existence of strongly contextual models can be used to study problems outside of quantum mechanics, including when a set of sentences is paradoxical.

Contents

1. Introduction

In the early 20th century, it became clear that classical mechanics was unable to account for a wealth of observations related to the behavior of electrons within the atom. Attempts to accommodate these discrepancies led to the development of quantum mechanics to treat such microscopic phenomena. Although the two theories are consistent in the sense that the classical theory holds in the macroscopic limit of the quantum theory, the starkly different mathematics of quantum mechanics poses a challenge to classical intuitions.

We will delve into the particular mathematical models in Section 2, but the key conceptual issue is the question of *realism*. Classical mechanics is a realist theory: given any physical system (e.g. electrons in an atom), all of its physical properties (e.g. position and momentum) have a definite value at any given time, regardless of the act of measurement. Einstein is reported as indicating that realism should hold as well for an atomic object, which is governed by quantum mechanics, as for the Moon, which is governed by classical mechanics [13].

According to quantum mechanics, however, this is not the case. An essential aspect of any mathematical formalism of quantum mechanics is the *Heisenberg* *uncertainty principle*, which states that there is a lower bound to the uncertainty of the position $x(t)$ and momentum $p(t)$ of a particle. This comes from the canonical commutation relation

$$[x, p] = i\hbar.$$

More generally, there is a lower bound of uncertainty for any physical quantities that do not commute. To this day, the meaning of the uncertainty which pervades quantum mechanics is up for debate. Einstein, Podolsky, and Rosen (EPR) [8] argued that uncertainty arises from our inability to access some hidden part of reality that gives rise to the part of reality we model with quantum mechanics. If that is true, then there is some *hidden variable theory* that reproduces the predictions of quantum mechanics.

In the decades following the publication of the EPR paper, physicists and mathematicians including David Bohm [4], John von Neumann, Greta Hermann, and J.S. Bell [3] published various results regarding what characteristics a hidden variable theory could or could not have. Bohm's work consists of the explicit construction of a particular hidden variable theory, whereas von Neumann, Hermann, and Bell sought generic constraints—or "no-go theorems"—for such theories. These efforts were subsumed in 1967 by the work of Simon Kochen and Ernst Specker [11].

The Kochen-Specker theorem demonstrates the impossibility of *non-contextual* hidden variable theories. *Contextuality* refers to the fact that the value of a physical quantity depends on the whole set of measurements—or the measurement context—to which a physical system is subjected. Consider the strange consequences of this condition: the measured value of any given physical quantity is affected by the other measurements you choose to make of the same system, even if those measurements occur later in time.

The pursuit of a complete characterization of contextuality remains an important problem in the mathematical foundations of quantum mechanics. Much progress has been made on this front since Mermin [12] showed that one could obtain simple proofs of contextuality using hypothetical scenarios later termed *all-versus-nothing* scenarios. A generalization of these scenarios due to Abramsky et al. [2] makes use of presheaf cohomology and is inspired by C.J. Isham's topos quantum theory [7]. In this paper, we survey these two developments, which study contextuality and quantum mechanics more broadly in a new mathematical setting—topos theory. We aim to provide a concise and well-motivated exposition that makes clear how the Kochen-Specker theorem inspires a new perspective on quantum theory and how that perspective, in turn, has something new to say about the Kochen-Specker theorem. Finally, we consider some examples of how contextuality arises outside of quantum mechanics and how this theory can be applied in such cases.

In Section 2, we give a brief account of hidden variable theories, contextuality, and the Kochen-Specker theorem. In Section 3, we discuss the foundations of topos quantum theory wherein arises a sheaf-theoretic Kochen-Specker theorem. In Section 4, we discuss how this inspires a more general sheaf-theoretic notion of strong contextuality. In Section 5, we discuss how strong contextuality is related to cohomological obstructions. In Sections 6 and 7, we take a step away from quantum mechanics, and look at strong contextuality as a more general phenomenon. In Section 6, we detail the relationship between strong contextuality and the theory of semantic paradox; in Section 7, we speculate on a relationship with the theory of directed graphs.

2. Kochen-Specker theorem

In classical mechanics, there is a state space Σ and a set of observables $\mathcal{O}$. The state space is the set of all possible states in which a physical system can be. An observable is a property of a physical system that we can measure. When we measure an observable, we associate to it a real number, so for any state $\psi \in \Sigma$, we have a valuation function $\lambda_\psi : \mathcal{O} \to \mathbb{R}$, where $\lambda_\psi(A)$ represents the value of observable A given state ψ. This means that classical mechanics is not only realist but *non-contextual*: the measurement context, or set of measurements actually performed on the system, is irrelevant to the value of $\lambda_\psi(A)$ for any given $A \in \mathcal{O}$. We define non-contextuality formally as follows.

Definition 2.1. A theory $(\Sigma, \mathcal{O})$ is *non-contextual* if every state $\psi \in \Sigma$ determines a valuation function $\lambda_\psi : \mathcal{O} \to \mathbb{R}$ subject to the *functional composition principle*: for any function $f : \mathbb{R} \to \mathbb{R}$ and observable $O \in \mathcal{O}$,

$$\lambda_\psi(f(O)) = f(\lambda_\psi(O)). \tag{2.2}$$

The functional composition principle ensures that if we, for example, measure the momentum of a system and then square that value, we get the same result as if we measure the squared momentum itself.

Definition 2.3. *Quantum mechanics* consists of

i) a state space $\mathcal{H}$, which is a Hilbert space with a Hermitian inner product $\langle -, - \rangle$;
ii) a set of observables $\mathcal{O}_{\text{sa}}$ consisting of bounded self-adjoint operators on $\mathcal{H}$.

The particular choice of Hilbert space depends on the physical system in question. For example, the spin of an electron is represented by a vector in the Hilbert space $\mathbb{C}^2$. In this paper, we consider finite-dimensional Hilbert spaces.

Remark 2.4. To understand (2.2) in this context, we need to clarify what we mean by $f(\hat{O})$ for a self-adjoint operator $\hat{O}$. If $\hat{O} = \sum o_i \hat{P}_i$ is the spectral decomposition of $\hat{O}$, then

$$f(\hat{O}) \coloneqq \sum f(o_i)\hat{P}_i.$$

Definition 2.5. A *valuation function* $\lambda : \mathcal{O}_{\text{sa}} \to \mathbb{R}$ for quantum mechanics maps any self-adjoint operator $\hat{O}$ to an element of its spectrum $\sigma(\hat{O})$ such that for all $f : \mathbb{R} \to \mathbb{R}$, the following holds

$$\lambda(f(\hat{O})) = f(\lambda(\hat{O})), \tag{2.6}$$

where $f(\hat{O})$ is defined as in Remark 2.4.

Kochen and Specker [11] set out to check whether one could actually construct such valuation functions for quantum mechanics. If it were possible to do so, then quantum mechanics would admit a refinement by a non-contextual hidden variable theory and could be considered a realist theory. Their efforts culminated in the following no-go theorem.

Theorem 2.7 (Kochen-Specker). *If* $\dim \mathcal{H} > 2$*, then there does not exist a valuation function* $\lambda : \mathcal{O}_{sa} \to \mathbb{R}$*.*

In order to prove this theorem, we need the following two results regarding operators and valuation functions.

Lemma 2.8. *Given $\hat{A}, \hat{B} \in \mathcal{O}_{sa}$ such that $[\hat{A}, \hat{B}] = \hat{0}$, there exist functions $f, g : \mathbb{R} \to \mathbb{R}$ and $\hat{C} \in \mathcal{O}_{sa}$ such that $\hat{A} = f(\hat{C})$ and $\hat{B} = g(\hat{C})$.*

Proof. Since $[\hat{A}, \hat{B}] = \hat{0}$, $\hat{A}$ and $\hat{B}$ can be simultaneously diagonalized, i.e. there is a set of projectors $\{\hat{P}_i\}$ such that $\hat{A} = \sum a_i \hat{P}_i$ and $\hat{B} = \sum b_i \hat{P}_i$. Then, we can choose some $\{c_i\} \subseteq \mathbb{R}$ and let $\hat{C} = \sum c_i \hat{P}_i$. Then, we can define $f, g : \mathbb{R} \to \mathbb{R}$ such that $f(c_i) = a_i$ and $g(c_i) = b_i$. Thus, $\hat{A} = f(\hat{C})$ and $\hat{B} = g(\hat{C})$. □

Lemma 2.9. *For $\hat{A}, \hat{B} \in \mathcal{O}_{sa}$ and valuation function λ,*

$$\lambda(\hat{A} + \hat{B}) = \lambda(\hat{A}) + \lambda(\hat{B}).$$

Proof. By Lemma 2.8, let $\hat{A} = f(\hat{C})$ and $\hat{B} = g(\hat{C})$, for some $\hat{C} \in \mathcal{O}_{sa}$. Define $h \coloneqq f + g$. Then,

$$\hat{A} + \hat{B} = f(\hat{C}) + g(\hat{C}) = h(\hat{C}).$$

By two applications of (2.6), we have

$$\lambda(\hat{A} + \hat{B}) = \lambda(h(\hat{C})) = h(\lambda(\hat{C})) = f(\lambda(\hat{C})) + g(\lambda(\hat{C})) = \lambda(f(\hat{C})) + \lambda(g(\hat{C})).$$

Thus, $\lambda(\hat{A} + \hat{B}) = \lambda(\hat{A}) + \lambda(\hat{B})$. □

Now we can discuss the proof of Theorem 2.7.

Proof. For ease of explanation, we will use a proof [5] that works only for $\dim \mathcal{H} > 3$. For a four-dimensional Hilbert space $\mathcal{H}$, we can find a set of 18 vectors such that we can construct 9 orthonormal bases (ONBs) of $\mathcal{H}$ by using each of our 18 vectors exactly twice. This is illustrated in the following diagram[1], in which each column defines an ONB. This makes it easy to check that each vector is used exactly twice.

e_1	(0,0,0,1)	(0,0,0,1)	(1,-1,1,-1)	(1,-1,1,-1)	(0,0,1,0)
e_2	(0,0,1,0)	(0,1,0,0)	(1,-1,-1,1)	(1,1,1,1)	(0,1,0,0)
e_3	(1,1,0,0)	(1,0,1,0)	(1,1,0,0)	(1,0,-1,0)	(1,0,0,1)
e_4	(1,-1,0,0)	(1,0,-1,0)	(0,0,1,1)	(0,1,0,-1)	(1,0,0,-1)

e_1	(1,-1,-1,1)	(1,1,-1,1)	(1,1,-1,1)	(1,1,1,-1)
e_2	(1,1,1,1)	(1,1,1,-1)	(-1,1,1,1)	(-1,1,1,1)
e_3	(1,0,0,-1)	(1,-1,0,0)	(1,0,1,0)	(1,0,0,1)
e_4	(0,1,-1,0)	(0,0,1,1)	(0,1,0,-1)	(0,1,-1,0)

For any ONB $\{e_1, e_2, e_3, e_4\}$, we can define a set of projectors by $\hat{P}_i \coloneqq e_i e_i^\dagger$, where $e_i^\dagger$ denotes the conjugate transpose of e_i. Then, we can consider a valuation function λ on these projectors.

$$\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \hat{P}_4 = \hat{\mathbb{1}}$$

$$\lambda(\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \hat{P}_4) = \lambda(\hat{\mathbb{1}})$$

Then by Lemma 2.9, $\lambda(\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \hat{P}_4) = \lambda(\hat{P}_1) + \lambda(\hat{P}_2) + \lambda(\hat{P}_3) + \lambda(\hat{P}_4)$. Moreover, $\lambda(\hat{\mathbb{1}})$ must be 1 since that is the only possible eigenvalue. Thus,

$$\lambda(\hat{P}_1) + \lambda(\hat{P}_2) + \lambda(\hat{P}_3) + \lambda(\hat{P}_4) = 1.$$

Projectors have eigenvalues of 0 and 1, so each $\lambda(\hat{P}_i)$ is either 0 or 1. Thus, exactly one such term is 1, and the rest are 0. Then, if we view λ as assigning a 0 or 1

[1]Reproduced from https://en.wikipedia.org/wiki/Kochen-Specker_theorem.

to every cell in the table above, there will be nine 1s, one for each ONB. However, each vector appears twice in the table, so there will be an even number of 1s. Thus, no such λ exists. □

Thus, quantum mechanics does not admit a refinement by a non-contextual hidden variable theory. A *contextual* hidden variable theory could still work, however. For example, if λ only had to be defined over one of the ONBs from the table above, then it would be easy to make an assignment consistent with (2.6). In general, if a valuation function λ_ψ depends on the context of measurements that the state ψ undergoes, then λ_ψ can be consistently defined. Put informally, a physical system must "know" what measurements to which it will be subject before it can assign values to observables.

Given that this condition is a difficult one to satisfy in a sensible way, why not just give up realism and hidden variable theories altogether? One possible answer is that belief in a realist interpretation of quantum mechanics is entailed by the belief that the universe is a closed system. After all, if you don't take a realist view of quantum mechanics, then you need an observer for any physical system to "become real." Thus, there must be an observer of the universe and also an observer of that observer and so on, *ad infinitum*.

So we do have an interest in studying hidden variable theories. How, then, should we go about constructing one? Since contextuality must be a feature of any such theory, we might as well bake it into the theory from the start, constructing a theory where observables depend on measurement contexts and there is no consistent way to "glue" all of these contexts together. This lead us to the mathematical framework of a topos of (pre)sheaves, which comes with a notion of gluing together local contexts. In the next section, we will look at how quantum mechanics can be studied within such a framework.

3. Topos quantum theory

Roughly speaking, topos quantum theory [7],[9] is formalized in the category of set-valued presheaves over a category of "measurement contexts," or sets of commuting observables that could be simultaneously measured of one system. We formalize this with the following definitions.

Definition 3.1. A *von Neumann algebra* is a $*$-algebra of bounded operators on a Hilbert space $\mathcal{H}$ that contains the identity and is closed in the weak operator topology.

Definition 3.2. Given $\mathcal{H}$, the category $\mathcal{V}(\mathcal{H})$ of *abelian von Neumann sub-algebras* is defined as follows. The objects of $\mathcal{V}(\mathcal{H})$ are abelian von Neumann algebras V of operators over $\mathcal{H}$, and for every $V, V' \in \mathcal{V}(\mathcal{H})$ such that $V \subseteq V'$, there exists a morphism $i_{V,V'} : V \to V'$.

The details of these definitions are not essential to understanding topos quantum theory or this paper, but we should note that $\mathcal{V}(\mathcal{H})$ is a category of *abelian* von Neumann algebras, so given some $V \in \mathcal{V}(\mathcal{H})$, every operator in V commutes. This means that every observable represented in a given $V \in \mathcal{V}(\mathcal{H})$ can be measured simultaneously, which makes V a classical context. The quantum behavior is encoded in the broader categorical structure of $\mathcal{V}(\mathcal{H})$ rather than within each $V \in \mathcal{V}(\mathcal{H})$.

Definition 3.3. Given $V \in \mathcal{V}(\mathcal{H})$, the *Gel'fand spectrum* Σ_V of V is the set of multiplicative linear functionals $\lambda : V \to \mathbb{C}$ such that $\lambda(\hat{\mathbb{1}}) = 1$ and for all $\hat{O} \in V$, $\lambda(\hat{O}) \in \sigma(\hat{O})$, where $\sigma(\hat{O})$ denotes the spectrum of $\hat{O}$.

We use the suggestive notation λ to imply that Σ_V is the set of valuation functions on V. This turns out to be the case.

Proposition 3.4. *Given a $V \in \mathcal{V}(\mathcal{H})$, the operators in V can be simultaneously diagonalized in some basis $\{e_i\}_{i\in I}$. Then, $\Sigma_V = \{\lambda_{e_i} : \hat{O} \mapsto \langle e_i, \hat{O}e_i\rangle\}_{i\in I}$.*

Proof. We first need to check that every λ_{e_i} is in Σ_V. By the definition of the Hermitian inner product, every λ_{e_i} is linear and satisfies $\lambda_{e_i}(\hat{\mathbb{1}}) = 1$, so we just need to check that λ_{e_i} is multiplicative:

$$\lambda_{e_i}(\hat{A}\hat{B}) = \langle e_i, \hat{A}\hat{B}e_i\rangle = \langle e_i, \hat{A}\lambda_{e_i}(\hat{B})e_i\rangle = \lambda_{e_i}(\hat{B})\,\langle e_i, \hat{A}e_i\rangle = \lambda_{e_i}(\hat{A})\lambda_{e_i}(\hat{B}).$$

Next, we need to show that any multiplicative linear functional $\lambda : V \to \mathbb{C}$ with $\lambda(\hat{\mathbb{1}}) = 1$ is a map of the form $\hat{O} \mapsto \langle e_i, \hat{O}e_i\rangle$. On account of linearity, we can characterize any given $\lambda \in \Sigma_V$ by its action on projectors $\{e_ie_i^\dagger\}_{i\in I}$. For $i \neq j$,

$$\lambda\left((e_ie_i^\dagger)(e_je_j^\dagger)\right) = \lambda(e_ie_i^\dagger e_je_j^\dagger) = \lambda(\hat{0}) = 0.$$

Thus by multiplicativity, either $\lambda(e_ie_i^\dagger) = 0$ or $\lambda(e_je_j^\dagger) = 0$. This means that λ is non-zero for at most one element of $\{e_ie_i^\dagger\}_{i\in I}$. The case where λ is zero for all elements of $\{e_ie_i^\dagger\}_{i\in I}$ violates the condition $\lambda(\hat{\mathbb{1}}) = 1$, so we suppose that for some $e_ie_i^\dagger$, we have $\lambda(e_ie_i^\dagger) \neq 0$. (Accordingly, for all $j \neq i$, we have $\lambda(e_je_j^\dagger) = 0$.) Then, by linearity, for any $\hat{A} = \sum_{j\in I} a_ie_je_j^\dagger$, we have $\lambda(\hat{A}) = a_i\lambda(e_ie_i^\dagger)$. Then,

$$\begin{aligned}\lambda(\hat{A})\lambda(e_ie_i^\dagger) = \lambda(\hat{A}e_ie_i^\dagger) &= \lambda\left(\sum_{j\in I} a_ie_je_j^\dagger e_ie_i^\dagger\right)\\ &= \lambda(a_ie_ie_i^\dagger) = a_i\lambda(e_ie_i^\dagger) = \lambda(\hat{A}).\end{aligned}$$

Thus, $\lambda(e_ie_i^\dagger) = 1$, so λ is the map $\hat{A} \mapsto \langle e_i, \hat{A}e_i\rangle$. □

Corollary 3.5. *Every $\lambda \in \Sigma_V$ satisfies (2.2) for all operators in V.*

Proof. Fix some $\lambda_{e_i} \in \Sigma_V$, $\hat{O} \in V$, and function $f : \mathbb{R} \to \mathbb{R}$. Note that we can write $\hat{O} = \sum_{j\in I} o_je_je_j^\dagger$. Then,

$$\lambda_{e_i}(f(\hat{O})) = \langle e_i, f(\hat{O})e_i\rangle = f(o_i) = f\left(\langle e_i, \hat{O}e_i\rangle\right) = f(\lambda_{e_i}(\hat{O})).$$

□

Corollary 3.6. *If $\lambda : \mathcal{O}_{sa} \to \mathbb{R}$ satisfies (2.2) for all self-adjoint operators in V, then $\lambda \in \Sigma_V$.*

Proof. Suppose $\lambda \notin \Sigma_V$. Then there exist $\hat{A}, \hat{B} \in V$ such that $\lambda(\hat{A}) = a_i$ and $\lambda(\hat{B}) = b_j$ for $i \neq j$, i.e. λ assigns two different operators to the eigenvalues associated with two different eigenvectors. Define $f : \mathbb{R} \to \mathbb{R}$ such that $f(\hat{A}) = \hat{B}$. Then, by (2.2),

$$\lambda(f(\hat{A})) = f(\lambda(\hat{A})) = f(a_i) = b_i,$$

but

$$\lambda(f(\hat{A})) = \lambda(\hat{B}) = b_j.$$

So we have a contradiction. Thus, $\lambda \in \Sigma_V$. □

Recall that in a realist theory valuation functions are associated to states, so we can consider Σ_V as the set of states in which a system can be if the system is measured in the context V. Then, we define the state space of topos quantum theory by putting all the Gel'fand spectra together with the following definition.

Definition 3.7. The *spectral presheaf* $\Sigma : \mathcal{V}(\mathcal{H})^{\mathrm{op}} \to \mathbf{Set}$ maps

i) every context V to its Gel'fand spectrum Σ_V;
ii) every morphism $i_{V'V} : V' \to V$ to the restriction map given by

$$\begin{aligned}\Sigma(i_{V'V}) : \Sigma(V) &\to \Sigma(V')\\ \lambda &\mapsto \lambda|_{V'}.\end{aligned}$$

Since Σ is the analogue of the state space, then there should be an analogous result to the Kochen-Specker theorem. To state this result, we need the following notion.

Definition 3.8. A *global section* of a presheaf $X : \mathscr{C}^{\mathrm{op}} \to \mathbf{Set}$ is a natural transformation $\gamma : 1 \to X$, where 1 denotes the terminal presheaf in $\mathbf{Set}^{\mathscr{C}^{\mathrm{op}}}$. whose components are given by elements $\gamma_A(\{*\}) \in X(A)$. Since γ is a natural transformation, this is subject to the condition that if there is some $f : B \to A$ in $\mathscr{C}$, then $X(f)(\gamma_A(\{*\})) = \gamma_B(\{*\})$.

Theorem 3.9. *The Kochen-Specker theorem is equivalent to the following: for* $\dim \mathcal{H} > 2$, *the spectral presheaf* Σ *has no global sections.*

Proof. We need to show that a global section of Σ exists if and only if a valuation function $\lambda : \mathcal{O}_{\mathrm{sa}} \to \mathbb{R}$ satisfying (2.2) exists.

Suppose that a global section γ of Σ exists. For any $\hat{O} \in \mathcal{O}_{\mathrm{sa}}$, there is some $V \in \mathcal{V}(\mathcal{H})$ such that $\hat{O} \in V$. Define λ by $\lambda(\hat{O}) \coloneqq \gamma_V(\hat{O})$. We need to check that this is well-defined. Consider some W such that $\hat{O} \in W$ but $V \neq W$. Then,

i) $V \cap W \subseteq V$, so $\gamma_{V\cap W}(\hat{O}) = \Sigma(i_{V\cap W,V})(\gamma_V)(\hat{O}) = \gamma_V(\hat{O})$, and
ii) $V \cap W \subseteq W$, so $\gamma_{V\cap W}(\hat{O}) = \Sigma(i_{V\cap W,W})(\gamma_W)(\hat{O}) = \gamma_W(\hat{O})$.

Thus, $\gamma_V(\hat{O}) = \gamma_W(\hat{O})$, so λ is well-defined. We need to check that λ satisfies (2.2). Fix some $V \in \mathcal{V}(\mathcal{H})$, $\hat{O} \in V$ and $f : \mathbb{R} \to \mathbb{R}$. Note that since $f(\hat{O})$ is defined via the spectrum of $\hat{O}$, $f(\hat{O})$ is either in V, in a superset of V, or in a subset of V—it commutes with every operator in V. Then, pick some $V' \supseteq V$ such that $\hat{O} \in V'$ and $f(\hat{O}) \in V'$. Then,

$$\lambda(f(\hat{O})) = \gamma_{V'}(f(\hat{O})) = f(\gamma_{V'}(\hat{O})) = f(\lambda(\hat{O})),$$

where the middle equality comes from Corollary 3.5. Thus, λ satisfies (2.2).

Alternatively, suppose that there exists a $\lambda : \mathcal{O}_{\mathrm{sa}} \to \mathbb{R}$ satisfying (2.2). Then by Corollary 3.6, for any $V \in \mathcal{V}(\mathcal{H})$, $\lambda|_V \in \Sigma_V$. Define a global section γ of Σ by $\gamma_V \coloneqq \lambda|_V$. We just need to check that this is a global section. Take some $V' \subseteq V$. Then,

$$\Sigma(i_{V',V})(\gamma_V) = \gamma_V|_{V'} = (\lambda|_V)|_{V'} = \lambda|_{V'} = \gamma_{V'}.$$

Thus, γ is a global section of Σ. □

On account of the existence of weaker forms of contextuality—see Bell's theorem [3] and Hardy states [10]—the kind of contextuality linked to the Kochen-Specker theorem is called *strong contextuality*. Theorem 3.9 suggests that strongly contextual scenarios can be studied in terms of global sections of certain presheaves. This suggestion inspires the following generic framework for studying strong contextuality.

4. Strong contextuality

A sheaf-theoretic characterization of strong contextuality in terms of strongly contextual empirical models is due to Abramsky et al. [2]; we summarize this work and then show that the existence of a certain strongly contextual empirical model suffices to prove the Kochen-Specker theorem.

Definition 4.1. A *measurement scenario* $(X, \mathcal{M}, O)$ consists of a finite set X of *measurements*, a finite set O of *outcomes*, and a *measurement cover* $\mathcal{M}$ of X consisting of *measurement contexts* $C \subseteq X$ such that

i) $\bigcup_{C\in\mathcal{M}} C = X$;
ii) if $C, C' \in \mathcal{M}$ and $C \subseteq C'$, then $C = C'$.

We denote the power set (the set of all subsets) of X by $\mathcal{P}(X)$ and we view $\mathcal{P}(X)$ as a poset ordered by subset inclusion.

Definition 4.2. Given a measurement scenario $(X, \mathcal{M}, O)$, the *event presheaf* $\mathcal{E} : \mathcal{P}(X)^{\mathrm{op}} \to \mathbf{Set}$ maps

i) every $U \subseteq X$ to the set $\mathcal{E}(U) \coloneqq \hom(U, O)$;
ii) every $i_{U',U} : U' \to U$ to the map
$$\mathcal{E}(i_{U',U}) : \mathcal{E}(U) \to \mathcal{E}(U')$$
$$t \mapsto t|_{U'}$$
where the restriction is just the restriction of the function $t : U \to O$ to $U' \subseteq U$.

Note that $\mathcal{E}$ is a sheaf if we give $\mathcal{P}(X)$ the discrete topology.

Definition 4.3. Given a measurement scenario $(X, \mathcal{M}, O)$, an *empirical model* $\mathcal{S}$ is a subpresheaf of $\mathcal{E}$ such that for any $C \in \mathcal{M}$, $\mathcal{S}(C) \neq \varnothing$ and such that $\mathcal{S}$ is flasque below the cover, i.e. for any $U' \subseteq U \subseteq C \in \mathcal{M}$, the map $\mathcal{S}(i_{U',U}) : \mathcal{S}(U) \to \mathcal{S}(U')$ is onto.

Definition 4.4. An empirical model $\mathcal{S}$ is *strongly contextual* (denoted $\mathsf{SC}(\mathcal{S})$) if it has no global sections.

Proposition 4.5. *A global section γ of $\mathcal{S}$ is an element of $\mathcal{S}(X)$.*

Proof. Fix a global section $\gamma : 1 \to \mathcal{S}$. Since $X \in \mathcal{P}(X)$, there is some component $\gamma_X \in \mathcal{S}(X)$. For any $U \subseteq X$ and any $x \in U$, we must have $\gamma_U(x) = \gamma_X(x)$, so all components of γ are defined by γ_X.

Alternatively, if there exists some $g \in \mathcal{S}(X)$, this defines a global section by $\gamma_U \coloneqq g|_U$. Thus, a global section of $\mathcal{S}$ is equivalent to an element of $\mathcal{S}(X)$. ☐

The following definition relates the data of the spectral presheaf Σ and an empirical model $\mathcal{S}$ in such a way that we can use strong contextuality to prove the Kochen-Specker theorem.

Definition 4.6. Given $\mathcal{H}$, a *quantum measurement scenario* is a measurement scenario $(X, \mathcal{M}, O)$ such that X is a finite subset of self-adjoint operators on $\mathcal{H}$ and O is the set of all possible outcomes of these operators, i.e.

$$O \coloneqq \bigcup_{\hat{O} \in X} \sigma(\hat{O}).$$

Theorem 4.7. *Given an empirical model $\mathcal{S}$ defined over a quantum measurement scenario, if $\mathcal{S}$ is strongly contextual, then the spectral presheaf Σ has no global sections.*

Proof. Suppose for sake of contradiction that there exists a global section γ of Σ. As shown in Theorem 3.9, this defines a function $\lambda : \mathcal{O}_{\text{sa}} \to \mathbb{R}$. We can define $g \in \mathcal{S}(X)$ by $g \coloneqq \lambda|_X$. Then, $\mathcal{S}(X)$ is non-empty, so we have a contradiction. Thus, there are no global sections of Σ. ∎

Theorem 4.7 links the sheaf-theoretic framework of strong contextuality with Theorem 3.9 and thus with the Kochen-Specker theorem. Moreover, this framework turns out to be the most useful for studying what kinds of empirical models are strongly contextual. While a complete characterization of strongly contextual models remains elusive, all known instances are accounted for and generalized by the notion of an *all-versus-nothing model*, which we examine in the next section. Moreover, the contextuality of every all-versus-nothing model is witnessed by a cohomological obstruction.

5. Cohomological witnesses of contextuality

5.1. **All-versus-nothing models.** All-versus-nothing models were introduced in Mermin [12] and generalized in Abramsky et al. [2]. Notably, this generalization is sufficient to account for every known example of strong contextuality. The generalized definition is inspired by the observation that every known strongly contextual model defines an unsolvable system of equations in some ring.

Fix a measurement scenario $(X, \mathcal{M}, R)$ where the set of outcomes is a finite commutative ring R.

Definition 5.1. An *R-linear equation* $\phi = \langle C, a, b\rangle$ consists of a context $C \in \mathcal{M}$, a function $a : C \to R$, and an element $b \in R$. We say that a section $s \in \mathcal{E}(C)$ *satisfies* ϕ (denoted $s \models \phi$) if

$$\sum_{m \in C} a(m)s(m) = b.$$

Let Γ be a set of R-linear equations. This in turn defines a set $\mathbb{M}(\Gamma)$ of sections that satisfy every equation in Γ, i.e.

$$\mathbb{M}(\Gamma) \coloneqq \{s \in \mathcal{E}(C) \mid \text{for all } \phi \in \Gamma,\ s \models \phi\}.$$

Alternatively, a set of sections $S \subseteq \mathcal{E}(C)$ defines an *R-linear theory*, which is a set $\mathbb{T}_R(S)$ of R-linear equations that are satisfied by all sections in S, i.e.

$$\mathbb{T}_R(S) \coloneqq \{\phi \mid \text{for all } s \in S \subseteq \mathcal{E}(C),\ s \models \phi\}.$$

Definition 5.2. Define the R-linear theory of an empirical model $\mathcal{S}$ over $(X, \mathcal{M}, R)$ by

$$\mathbb{T}_R(\mathcal{S}) \coloneqq \bigcup_{C \in \mathcal{M}} \mathbb{T}_R(\mathcal{S}(C)) = \bigcup_{C \in \mathcal{M}} \{\phi \mid \text{for all } s \in \mathcal{S}(C),\ s \models \phi\}.$$

Definition 5.3. $\mathcal{S}$ is an *all-versus-nothing model over* R (denoted $\mathsf{AvN}_R(\mathcal{S})$) if there is no function $g : X \to R$ such that for all $\phi = \langle C, a, b\rangle \in \mathbb{T}_R(\mathcal{S})$, $g|_C$ satisfies ϕ.

Proposition 5.4. $\mathsf{AvN}_R(\mathcal{S}) \Rightarrow \mathsf{SC}(\mathcal{S})$ *(see Definition 4.4))*

Proof. For sake of contradiction, suppose that $\mathcal{S}$ is not strongly contextual, so there exists a global section $g : X \to R$. Let $\phi \in \mathbb{T}_R(\mathcal{S})$. Then, for some $C \in \mathcal{M}$, $\phi \in \mathbb{T}_R(\mathcal{S}(C))$. Note that $g|_C \in \mathcal{S}(C)$, so by Definition 5.2, $g|_C$ satisfies ϕ. This argument works for all $\phi \in \mathbb{T}_R(\mathcal{S})$, so $\mathcal{S}$ is not an all-versus-nothing model. This is a contradiction, so $\mathsf{SC}(\mathcal{S})$. □

Definition 5.5. The *affine closure* $\operatorname{aff} S$ of a set $S \subseteq \mathcal{E}(U)$ is given by

$$\operatorname{aff} S \coloneqq \left\{ \sum_{i=1}^{t} c_i s_i \mid s_i \in S, c_i \in R, \sum_{i=1}^{t} c_i = 1 \right\}.$$

Definition 5.6. The *affine closure of an empirical model* $\mathcal{S}$ is defined for each $C \in \mathcal{M}$ by

$$(\operatorname{Aff}\mathcal{S})(C) \coloneqq \operatorname{aff}(\mathcal{S}(C)).$$

Proposition 5.7. $\mathsf{AvN}_R(\mathcal{S}) \Rightarrow \mathsf{AvN}_R(\operatorname{Aff}\mathcal{S}) \Rightarrow \mathsf{SC}(\operatorname{Aff}\mathcal{S})$

Proof. Note that $(\mathbb{M} \circ \mathbb{T})(S)$ gives the maximum set of sections that satisfy every equation satisfied by all sections in S. Since the affine combination of solutions to a linear equation is also a solution, $\operatorname{aff} S \subseteq (\mathbb{M} \circ \mathbb{T})(S)$. Thus, for all $\mathcal{S}(C)$, $\mathbb{T}_R(\operatorname{aff}(\mathcal{S}(C))) = \mathbb{T}_R(\mathcal{S}(C))$. Thus, $\mathbb{T}_R(\mathcal{S}) = \mathbb{T}_R(\operatorname{Aff}\mathcal{S})$, so if $\mathcal{S}$ is an AvN_R model, then so is $\operatorname{Aff}\mathcal{S}$. Then by Proposition 5.4, $\operatorname{Aff}\mathcal{S}$ is strongly contextual. □

The reason for shifting to the affine model is to show that the strong contextuality of all-versus-nothing models is witnessed by cohomological obstructions.

5.2. **Čech cohomology.** We will briefly work in a more general setting than above in order to give an account of the Čech cohomology theory. Fix a topological space X, open cover $\mathcal{M}$, and presheaf $\mathcal{F} : \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{AbGrp}$.

Definition 5.8. A *q-simplex of the nerve of* $\mathcal{M}$ is a $(q+1)$-tuple $\sigma = (C_0, \ldots, C_q)$ of elements of $\mathcal{M}$ such that

$$|\sigma| \coloneqq \bigcap_{C \in \sigma} C \neq \varnothing.$$

Let $\mathcal{N}(\mathcal{M})_q$ denote the set of q-simplices.

Definition 5.9. Define the *j-th partial boundary operator* ∂_j by

$$\begin{aligned} \partial_j : \mathcal{N}(\mathcal{M})_q &\to \mathcal{N}(\mathcal{M})_{q-1} \\ (C_0, \ldots, C_j, \ldots, C_q) &\mapsto (C_0, \ldots, C_{j-1}, C_{j+1}, \ldots, C_q). \end{aligned}$$

Note that $|\sigma| \subseteq |\partial_j \sigma|$, so there exists a morphism $i_{|\sigma|,|\partial_j\sigma|} : |\sigma| \to |\partial_j\sigma|$ in $\mathcal{O}(X)$. Thus, there exists a restriction map

$$\rho^{|\partial_j\sigma|}_{|\sigma|} \coloneqq \mathcal{F}(i_{|\sigma|,|\partial_j\sigma|}) : \mathcal{F}(|\partial_j\sigma|) \to \mathcal{F}(|\sigma|).$$

Definition 5.10. A *q-cochain* ω is a map on q-simplices such that for all $\sigma \in \mathcal{N}(\mathcal{M})_q$, $\omega(\sigma) \in \mathcal{F}(|\sigma|)$. The set $C^q(\mathcal{M}, \mathcal{F})$ of q-cochains forms an abelian group under pointwise addition. That is, if $\omega_1, \omega_2 \in C^q(\mathcal{M}, \mathcal{F})$, then so is $\omega_1 + \omega_2$, which is defined by $(\omega_1 + \omega_2)(\sigma) \coloneqq \omega_1(\sigma) + \omega_2(\sigma)$. Since $\omega_1(\sigma)$ and $\omega_2(\sigma)$ are elements of the same abelian group $\mathcal{F}(|\sigma|)$, then $(\omega_1 + \omega_2)(\sigma) \in \mathcal{F}(|\sigma|)$, so $\omega_1 + \omega_2 \in C^q(\mathcal{M}, \mathcal{F})$.

Definition 5.11. Define the *coboundary operator* $\delta_q : C^q(\mathcal{M}, \mathcal{F}) \to C^{q+1}(\mathcal{M}, \mathcal{F})$ by

$$(\delta_q \omega)(\sigma) \coloneqq \sum_{j=0}^{q+1} (-1)^j \rho^{|\partial_j \sigma|}_{|\sigma|}(\omega(\partial_j \sigma)).$$

Definition 5.12. A *q-cocycle* is a q-cochain ω such that for all $(q+1)$-simplices σ, $(\delta_q \omega)(\sigma) = 0$. Denote the set of q-cocycles as $Z^q(\mathcal{M}, \mathcal{F}) \coloneqq \ker \delta_q$. A *q-coboundary* is a q-cochain ω such that there exists a $(q-1)$-cochain τ such that $\delta_{q-1} \tau = \omega$. Denote the set of q-coboundaries as $B^q(\mathcal{M}, \mathcal{F}) \coloneqq \operatorname{Im} \delta_{q-1}$.

Since $\delta_{q+1} \circ \delta_q = 0$, we have $B^q(\mathcal{M}, \mathcal{F}) \subseteq Z^q(\mathcal{M}, \mathcal{F}) \subseteq C^q(\mathcal{M}, \mathcal{F})$.

Definition 5.13. We have the following *cochain complex* $(C^\bullet(\mathcal{M}, \mathcal{F}), \delta)$:

$$0 \to C^0(\mathcal{M}, \mathcal{F}) \xrightarrow{\delta_0} C^1(\mathcal{M}, \mathcal{F}) \xrightarrow{\delta_1} C^2(\mathcal{M}, \mathcal{F}) \xrightarrow{\delta_2} \cdots$$

The *Čech cohomology* of $\mathcal{M}$ in $\mathcal{F}$ is the cohomology of $(C^\bullet(\mathcal{M}, \mathcal{F}), \delta)$, i.e. the q-th Čech cohomological group is $\check{H}^q(\mathcal{M}, \mathcal{F}) \coloneqq Z^q(\mathcal{M}, \mathcal{F}) / B^q(\mathcal{M}, \mathcal{F})$.

Remark 5.14. Note that a 0-simplex is an element $C \in \mathcal{M}$. Thus, a 0-cochain is a family $\{r_C \in \mathcal{F}(C)\}_{C \in \mathcal{M}}$ and a 0-coycle is a compatible family[2]. Moreover, $\check{H}^0(\mathcal{M}, \mathcal{F}) \coloneqq Z^0(\mathcal{M}, \mathcal{F})/0 \cong Z^0(\mathcal{M}, \mathcal{F})$, so $\check{H}^0(\mathcal{M}, \mathcal{F})$ is the set of compatible families.

Remark 5.15. Since $B^1(\mathcal{M}, \mathcal{F}) \subseteq Z^1(\mathcal{M}, \mathcal{F})$, we can corestrict δ_0 to a map $\tilde{\delta}_0 : C^0(\mathcal{M}, \mathcal{F}) \to Z^1(\mathcal{M}, \mathcal{F})$. Then, $\ker \tilde{\delta}_0 = Z^0(\mathcal{M}, \mathcal{F}) \cong \check{H}^0(\mathcal{M}, \mathcal{F})$ and $\operatorname{coker} \tilde{\delta}_0 = Z^1(\mathcal{M}, \mathcal{F})/B^1(\mathcal{M}, \mathcal{F}) = \check{H}^1(\mathcal{M}, \mathcal{F})$.

Now we fix an open set $U \subset X$.

Definition 5.16. Let $\mathcal{F}|_U : \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{AbGrp}$ be defined by $\mathcal{F}|_U(V) \coloneqq \mathcal{F}(U \cap V)$. There is a natural transformation $p : \mathcal{F} \to \mathcal{F}|_U$ whose components are given by

$$\begin{aligned} p_V : \mathcal{F}(V) &\to \mathcal{F}(U \cap V) \\ r &\mapsto r|_{U \cap V}. \end{aligned}$$

Let $\mathcal{F}_{\tilde{U}} : \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{AbGrp}$ be defined by $\mathcal{F}_{\tilde{U}}(V) \coloneqq \ker p_V \subseteq \mathcal{F}(V)$.

Thus, we have the following exact sequence:

$$0 \to \mathcal{F}_{\tilde{U}} \hookrightarrow \mathcal{F} \xrightarrow{p} \mathcal{F}|_U \tag{5.17}$$

Moreover, we can lift this to a short exact sequence at the level of cochain groups:

$$0 \to C^0(\mathcal{M}, \mathcal{F}_{\tilde{U}}) \to C^0(\mathcal{M}, \mathcal{F}) \to C^0(\mathcal{M}, \mathcal{F}|_U) \to 0 \tag{5.18}$$

[2]If $\omega = \{r_C \in \mathcal{F}(C)\}_{C \in \mathcal{M}}$ is a 0-coycle, then for any 1-simplex (C_0, C_1), $(\delta_0 \omega)((C_0, C_1)) = 0$. That is, $\rho^{C_1}_{C_0 \cap C_1}(\omega(C_1)) - \rho^{C_0}_{C_0 \cap C_1}(\omega(C_0)) = 0$, or $\rho^{C_1}_{C_0 \cap C_1}(r_{C_1}) = \rho^{C_0}_{C_0 \cap C_1}(r_{C_0})$.

Putting the map from Remark 5.15 and (5.18) together, we have the diagram

$$\begin{array}{ccccccccc}
 & & \check{H}^0(\mathcal{M},\mathcal{F}_{\tilde{U}}) & \longrightarrow & \check{H}^0(\mathcal{M},\mathcal{F}) & \longrightarrow & \check{H}^0(\mathcal{M},\mathcal{F}|_U) & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & C^0(\mathcal{M},\mathcal{F}_{\tilde{U}}) & \longrightarrow & C^0(\mathcal{M},\mathcal{F}) & \longrightarrow & C^0(\mathcal{M},\mathcal{F}|_U) & \longrightarrow & 0 \\
 & & \downarrow{\tilde{\delta}_0} & & \downarrow{\tilde{\delta}_0} & & \downarrow{\tilde{\delta}_0} & & \\
0 & \longrightarrow & Z^1(\mathcal{M},\mathcal{F}_{\tilde{U}}) & \longrightarrow & Z^1(\mathcal{M},\mathcal{F}) & \longrightarrow & Z^1(\mathcal{M},\mathcal{F}|_U) & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{U}}) & \longrightarrow & \check{H}^1(\mathcal{M},\mathcal{F}) & \longrightarrow & \check{H}^1(\mathcal{M},\mathcal{F}|_U) & &
\end{array}$$

where the groups in the first row are the kernels of the map between the second and third row and the groups in the last row are the cokernels. Thus, by the snake lemma there exists a map $\gamma : \check{H}^0(\mathcal{M},\mathcal{F}|_U) \to \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{U}})$ such that

$$\check{H}^0(\mathcal{M},\mathcal{F}_{\tilde{U}}) \to \check{H}^0(\mathcal{M},\mathcal{F}) \to \check{H}^0(\mathcal{M},\mathcal{F}|_U) \xrightarrow{\gamma} \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{U}}) \to \check{H}^1(\mathcal{M},\mathcal{F}) \to \check{H}^1(\mathcal{M},\mathcal{F}|_U)$$

is an exact sequence.

Remark 5.19. Recall from Remark 5.14 that $\check{H}^0(\mathcal{M},\mathcal{F}|_U)$ is the set of compatible families of sections in $\mathcal{F}|_U$. Consider such a family $\{s_C \in \mathcal{F}|_U(C)\}_{C\in\mathcal{M}}$. By Definition 5.16, this is equal to $\{s_C \in \mathcal{F}(U \cap C)\}_{C\in\mathcal{M}}$, and this family uniquely determines a section $s \in \mathcal{F}(U)$ such that for all $C \in \mathcal{M}$, $s|_{U\cap C} = s_C$. Thus, $\check{H}^0(\mathcal{M},\mathcal{F}|_U) \cong \mathcal{F}(U)$.

Remark 5.20. Consider the map $\check{H}^0(\mathcal{M},\mathcal{F}) \to \check{H}^0(\mathcal{M},\mathcal{F}|_{C_0})$ where $C_0 \in \mathcal{M}$. Call this map P. An element $\omega \in \check{H}^0(\mathcal{M},\mathcal{F})$ is a compatible family $\omega = \{r_C \in \mathcal{F}(C)\}_{C\in\mathcal{M}}$. Then,

$$P(\omega) = \{p_C(r_C) \in \mathcal{F}(C_0 \cap C)\}_{C\in\mathcal{M}}$$

where each p_C is a component of the natural transformation $p : \mathcal{F} \to \mathcal{F}|_{C_0}$. The element r_{C_0} of the original family associated with C_0 is unchanged. That is, $r_{C_0} \in P(\omega)$. Moreover, since $r_{C_0} \in \mathcal{F}(C_0)$, it must be that r_{C_0} is the unique section determined by the compatible family $P(\omega)$ as discussed in Remark 5.19. Thus, P maps a compatible family $\{r_C \in \mathcal{F}(C)\}_{C\in\mathcal{M}}$ to the section r_{C_0} of the family associated to $C_0 \in \mathcal{M}$.

Definition 5.21. Given $C_0 \in \mathcal{M}$ and $r_0 \in \mathcal{F}(C_0)$, the *cohomological obstruction* of r_0 is $\gamma(r_0)$, where $\gamma : \check{H}^0(\mathcal{M},\mathcal{F}|_{C_0}) \to \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{C}_0})$.

Proposition 5.22. *$\gamma(r_0) = 0$ if and only if there exists a compatible family $\{r_C \in \mathcal{F}(C)\}_{C\in\mathcal{M}}$ such that $r_{C_0} = r_0$.*

Proof. From Remarks 5.19 and 5.20, we have the following exact sequence:

$$\check{H}^0(\mathcal{M},\mathcal{F}) \xrightarrow{P} \mathcal{F}(C_0) \xrightarrow{\gamma} \check{H}^1(\mathcal{M},\mathcal{F})$$

Thus, $\gamma(r_0) = 0$ if and only if $r_0 \in \ker\gamma = \operatorname{Im} P$. Then by the definition of P, $r_0 \in \operatorname{Im} P$ if and only if there exists a compatible family $\{r_C \in \mathcal{F}(C)\}_{C\in\mathcal{M}}$. □

So we are justified in referring to $\gamma(r_0)$ as an "obstruction." It tells us whether or not we can extend a particular local section to a global section.

5.3. **Witnesses for AvN models.** In order to find cohomological obstructions for empirical models, we need to make one modification. In particular, we need to turn empirical models into functors to abelian groups rather than sets.

Definition 5.23. Fix a commutative ring R and define the functor $F_R : \mathbf{Set} \to \mathbf{Mod}_R$ to the category of modules over R as the functor that maps

i) every set X to the free R-module $F_R(X) = \{\phi : X \to R \mid \operatorname{supp}\phi < \infty\}$;
ii) every function $f : X \to Y$ to the map

$$F_R(f) : F_R(X) \to F_R(Y)$$
$$\phi \mapsto \lambda,$$

where

$$\lambda(y) = \sum_{f(x)=y} \phi(x).$$

Definition 5.24. An empirical model $\mathcal{S}$ over a measurement scenario $(X, \mathcal{M}, O)$ is *cohomologically strongly contextual over a ring* R (denoted $\mathsf{CSC}_R(\mathcal{S})$) if for all $C \in \mathcal{M}, s \in C$, we have the obstruction $\gamma_{F_R\mathcal{S}}(s) \neq 0$.

Note that we can view an element $\phi \in F_R(X)$ as a formal linear combination of elements of X given by $\sum \phi(x)x$.

Proposition 5.25. $\mathsf{CSC}_R(\mathcal{S}) \Rightarrow \mathsf{SC}(\mathcal{S})$

Proof. By Proposition 5.22, for all $C \in \mathcal{M}, s \in C$, s is not a member of any compatible family. Thus, there is no compatible family over $\mathcal{M}$, so there is no global section of $\mathcal{S}$. Thus, $\mathcal{S}$ is strongly contextual. □

Since AvN models encompass all known strongly contextual models, we would like to show that they are cohomologically strongly contextual. From Proposition 5.7, we have $\mathsf{AvN}_R(\mathcal{S}) \Rightarrow \mathsf{SC}(\mathrm{Aff}\,\mathcal{S})$. Thus, we need to show $\mathsf{SC}(\mathrm{Aff}\,\mathcal{S}) \Rightarrow \mathsf{CSC}_R(\mathcal{S})$.

Remark 5.26. Since F_R maps sets to the free R-module over that set, it forms an adjoint pair with the forgetful functor $U : \mathbf{Mod}_R \to \mathbf{Set}$. The counit of the adjunction $\epsilon : F_R U \to 1_{\mathbf{Mod}_R}$ is given by components ϵ_M defined by

$$\epsilon_M : F_R U(M) \to M$$
$$r \mapsto \sum_{m \in M} r(m)m.$$

Since the set of outcomes is the ring R, every $\mathcal{E}(C)$ is itself an R-module. Then, we have the map $\epsilon : F_R U\mathcal{E} \to \mathcal{E}$ given by $\epsilon_{\mathcal{E}(C)} : F_R U\mathcal{E}(C) \to \mathcal{E}(C)$. For $\mathcal{S} \subseteq \mathcal{E}$, this restricts to a map $F_R U\mathcal{S} \to \operatorname{Span}\mathcal{S}$ given by $\epsilon_{\mathcal{S}(C)} : F_R U\mathcal{S}(C) \to \operatorname{Span}\mathcal{S}(C)$, where $\operatorname{Span}\mathcal{S}(C)$ denotes the linear span of $\mathcal{S}(C)$ in R. Finally, if we restrict to affine R-modules, we have a map $F_R^{\mathrm{aff}} U\mathcal{S} \to \mathrm{Aff}\,\mathcal{S}$.

Proposition 5.27. $\mathsf{SC}(\mathrm{Aff}\,\mathcal{S}) \Rightarrow \mathsf{CSC}_R(\mathcal{S})$

Proof. Suppose for sake of contradiction that there is some $C_0 \in \mathcal{M}$ such that there is some $s_0 \in \mathcal{S}(C_0)$ such that $\gamma_{F_R\mathcal{S}}(s_0) \neq 0$. By Proposition 5.22, there is a compatible family $\{r_C \in F_R\mathcal{S}(C)\}_{C \in \mathcal{M}}$ such that $r_{C_0} = s_0$. s_0 is not only a formal linear combination but an affine combination, since $s_0 = 1 \cdot s_0$.

Note that $\mathcal{S}(\varnothing) = \{*\}$. Then, for any C, we have $i : \varnothing \to C$, so we have the restriction

$$F_R i : F_R\mathcal{S}(C) \to F_R\mathcal{S}(\varnothing) = F_R\{*\}$$
$$r_C \mapsto r_C|_\varnothing,$$

where by Definition 5.23,

$$r_C|_\varnothing(*) = \sum_{s\in\mathcal{S}(C)} r_C(s).$$

Since we have a compatible family, every restriction to the empty set must be the same, so so the coefficients of every r_C sum to the same number. Then, since $r_{C_0} = s_0$ is an affine combination, so is every r_C. Then, the map $F_R^{\mathrm{aff}} U\mathcal{S} \to \mathrm{Aff}\mathcal{S}$ lifts the compatible family to a compatible family $\{r_C \in (\mathrm{Aff}\mathcal{S})(C)\}_{C\in\mathcal{M}}$, which contradicts $\mathsf{SC}(\mathrm{Aff}\mathcal{S})$. Thus, $\mathcal{S}$ is cohomologically strongly contextual. □

To summarize, we have the following for an empirical model $\mathcal{S}$:

$$\mathsf{AvN}_R(\mathcal{S}) \Rightarrow \mathsf{SC}(\mathrm{Aff}\mathcal{S}) \Rightarrow \mathsf{CSC}_R(\mathcal{S}) \Rightarrow \mathsf{SC}(\mathcal{S}).$$

All-versus-nothing models provide a good account of strong contextuality in quantum mechanics, and this is their intended purpose. However, there is sufficient generality in the idea to think about other ways strong contextuality emerges in mathematics. We now briefly discuss two examples where strong contextuality provides a nice perspective. The first example is the theory of semantic paradoxes.

6. Semantic paradox

The results of the previous section were proven in Abramsky et al. [2], where the authors note the possibility of applying their work to the case of semantic paradoxes. In this section, we demonstrate one possible approach that connects the contextuality framework outlined above to the characterization of semantic paradoxes in Cook [6].

6.1. Background.

A semantic paradox is a set of sentences such that truth values cannot be consistently assigned. A simple example is the *liar paradox*: "This sentence is false." If the liar paradox is false, then is is true, and if the liar paradox is true, then it is false. More generally, we can come up with large and even infinite sets of sentences that are paradoxical. The following characterization of semantic paradoxes in terms of a simple propositional language is from Cook [6].

Definition 6.1. We define our propositional language L_p as consisting of a class of sentence names $\{S_n\}_{n\in\mathbb{N}}$, a conjunction operation $\wedge$, and a negation $\neg$. Associated to L_p is a set of *well-formed formulae*, denoted WFF, given by

$$WFF := \{\wedge\{\neg(S_\beta)\}_{\beta\in B} \mid B \subseteq \mathbb{N}\}.$$

It is helpful to also use the following notation for a formula.

$$\neg(S_{\beta_1}) \wedge \neg(S_{\beta_2}) \wedge \neg(S_{\beta_3}) \wedge \cdots := \wedge\{\neg(S_\beta)\}_{\beta\in B}$$

The intuitive way to understand this notation is to take $\neg(S_\beta)$ to mean "the sentence S_β is false" and to take $\wedge$ to mean "and."

Definition 6.2. A *denotation function* δ is a map

$$\delta : \{S_n\}_{n\in\mathbb{N}} \to WFF.$$

δ assigns each sentence to a particular "meaning" in the form of a well-formed formula. For each S_n, we define the set

$$D_\delta(S_n) := \{S_\gamma \mid \delta(S_n) = \wedge\{\neg(S_\beta)\}_{\beta\in B} \text{ and } \gamma \in B\}.$$

Intuitively, we can think of $D_\delta(S_n)$ as the set of sentences that must be false in order for S_n to be true. We formalize this idea with the following notion.

Definition 6.3. Let $T = \{0,1\}$. A *truth-value assignment* $\sigma : \{S_\beta\}_{\beta\in B} \to T$ is *acceptable* if for all $\beta \in B$, $\sigma(S_\beta) = 1$ if and only if for all $S_\gamma \in D_\delta(S_\beta)$, $\sigma(S_\gamma) = 0$.

We now have a formal notion of paradox.

Definition 6.4.

i) $\{S_\beta\}_{\beta\in B}$ is *evaluable* if there exists an acceptable assignment σ;
ii) $\{S_\beta\}_{\beta\in B}$ is *paradoxical* if there is no acceptable assignment.

We are interested in the following two situations.

Definition 6.5. The *n-cycle* consists of the set of sentences $\{S_i\}_{i\leq n}$ and a denotation function defined by

$$\delta(S_i) = \begin{cases} \neg(S_{i+1}) & i < n \\ \neg(S_1) & i = n \end{cases}$$

Definition 6.6. The *n-Yablo chain* consists of the set of sentences $\{S_i\}_{i\in\mathbb{N}}$ with the denotation function δ given by

$$\delta(S_i) = \wedge\{\neg(S_k) \mid \text{there exists } m \geq 0 \text{ such that } k = i + nm + 1\}.$$

Example 6.7. The 1-Yablo chain is known as *Yablo's paradox* [15]. The denotation of each sentence is

$$\begin{aligned} \delta(S_1) &= \neg(S_2) \wedge \neg(S_3) \wedge \neg(S_4) \wedge \cdots \\ \delta(S_2) &= \neg(S_3) \wedge \neg(S_4) \wedge \neg(S_5) \wedge \cdots \\ &\vdots \\ \delta(S_n) &= \neg(S_{n+1}) \wedge \neg(S_{n+2}) \wedge \neg(S_{n+3}) \wedge \cdots \\ &\vdots \end{aligned}$$

Yablo's paradox is a countable sequence of sentences, each one saying "all of the following sentences are false." Similarly, each sentence of the n-Yablo chain says "every n-th sentence of the following sentences is false."

There is an important connection between n-cycles and n-Yablo chains. To understand it, we need to understand the notion of "unwinding" a set of sentences.

Definition 6.8. Let $\prec$ denote the *lexicographical order* on pairs of natural numbers (x, y), i.e. $(a, b) \prec (c, d)$ if one of the following is true.

i) $a < c$
ii) $a = c$ and $b < d$

Definition 6.9. The *unwinding* of a finite set of sentences $\{S_i\}_{i\leq n}$ with denotation function δ is given by

$$(\{S_i\}_{i\leq n})^U = \{S_{(a,b)} \mid a, b \in \mathbb{N} \text{ and } b \leq n\}$$

with denotation function δ^U defined by

$$\delta^U(S_{(a,b)}) = \wedge\{\neg(S_{(c,d)}) \mid (a,b) \prec (c,d) \text{ and } S_d \in D_\delta(S_b)\}.$$

Definition 6.10. An assignment σ on $(\{S_i\}_{i\leq n})^U$ is *recurrent* if for all $a < b$ and $c < n$, $\sigma(S_{(a,c)}) = \sigma(S_{(b,c)})$.

Lemma 6.11. *If σ is acceptable on $(\{S_i\}_{i\leq n})^U$, then σ is recurrent.*

Proof. Let $a < b$ and $c < n$.

i) If $S_{(a,c)} = 1$, then for any (x,y) such that $(b,c) \prec (x,y)$, and $S_y \in D_\delta(S_c)$, we have $(a,c) \prec (x,y)$, so $\sigma(S_{(x,y)}) = 0$. But $S_{(x,y)} \in D_{\delta^U}(S_{(b,c)})$, so $S_{(b,c)} = 1 = S_{(a,c)}$.
ii) If $S_{(a,c)} = 0$, then there is some (j,k) such that $(a,c) \prec (j,k)$ and $S_k \in D_\delta(S_c)$ and $\sigma(S_{(j,k)}) = 1$. Suppose for sake of contradiction that $\sigma(S_{(b,c)}) = 1$. Then, since $(b,c) \prec (b+j+1,k)$ and $S_k \in D_\delta(S_c)$, $\sigma(S_{(b+j+1,k)}) = 0$. But $(j,k) \prec (b+j+1,k)$, so $S_{(b+j+1,k)} \in D_{\delta^U}(S_{(a,c)})$, so $S_{(a,c)} = 1$. This is a contradiction, so we must have $S_{(b,c)} = 0$.

Thus, σ is recurrent. □

Theorem 6.12. *σ is acceptable on $\{S_i\}_{i\leq n}$ if and only if σ' is acceptable on $(\{S_i\}_{i\leq n})^U$, where*

$$\sigma'(S_{(a,c)}) \coloneqq \sigma(S_c).$$

Proof. Suppose that σ is acceptable on $\{S_i\}_{i\leq n}$.

i) If $\sigma'(S_{(b,c)}) = 1$, then $\sigma(S_c) = 1$, so for all $S_d \in D_\delta(S_c)$, $\sigma(S_d) = 0$. Then, for all $S_{(a,d)} \in D_{\delta^U}(S_{(b,c)})$, $\sigma'(S_{(a,d)}) = 0$.
ii) If $\sigma'(S_{(b,c)}) = 0$, then $\sigma(S_c) = 0$, so there is some $S_d \in D_\delta(S_c)$ such that $S_d = 1$. Thus, there is some $a > b$ such that $S_{(a,d)} = 1$ and $S_{(a,d)} \in D_{\delta^U}(S_{(b,c)})$.

Thus, σ' is acceptable.

Alternatively, suppose that σ' is acceptable on $(\{S_i\}_{i\leq n})^U$. By Lemma 6.11, σ' is recurrent, i.e. for all $a < b$, $\sigma(S_{(a,c)}) = \sigma(S_{(b,c)})$. For all d, c such that $S_d \in D_\delta(S_c)$ and all $a < b$, we have $S_{(b,d)} \in D_{\delta^U}(S_{(a,c)})$. Then,

i) if $\sigma(S_c) = 1$, then $\sigma'(S_{(a,c)}) = 1$, so $\sigma'(S_{(b,d)}) = 0$, so $\sigma(S_d) = 0$;
ii) if $\sigma(S_c) = 0$, then $\sigma'(S_{(a,c)}) = 0$. This is true for all c such that $S_d \in D_\delta(S_c)$ and all $a < b$, so it holds for all $S_{(a,c)} \in D_{\delta^U}(S_{(b,d)})$. Thus, $S_{(b,d)} = 1$, so $S_d = 1$. This holds for all $S_d \in D_\delta(S_c)$.

Thus, σ is acceptable. □

Corollary 6.13. *$(\{S_i\}_{i\leq n})^U$ is paradoxical if and only if $\{S_i\}_{i\leq n}$ is paradoxical.*

We can apply this result to n-cycles and n-Yablo chains due to the following theorem.

Theorem 6.14. *The unwinding of the n-cycle is the n-Yablo chain.*

Proof. Consider the n-cycle, consisting of the set $\{S_i\}_{i\leq n}$ and denotation function δ as defined in Definition 6.5. Then, consider the unwinding

$$(\{S_i\}_{i\leq n})^U = \{S_{(a,1)} \mid a \in \mathbb{N}\} \cup \{S_{(a,2)} \mid a \in \mathbb{N}\} \cup \cdots \cup \{S_{(a,n)} \mid a \in \mathbb{N}\}.$$

The associated denotation function δ^U is given on each of the above subsets by

$$\begin{array}{rcl} \delta^U(S_{(a,1)}) & = & \wedge\{\neg(S_{(c,2)}) \mid a \leq c\} \\ \delta^U(S_{(a,2)}) & = & \wedge\{\neg(S_{(c,3)}) \mid a \leq c\} \\ \vdots & & \vdots \\ \delta^U(S_{(a,n-1)}) & = & \wedge\{\neg(S_{(c,n)}) \mid a \leq c\} \\ \delta^U(S_{(a,n)}) & = & \wedge\{\neg(S_{(c,1)}) \mid a < c\} \end{array}$$

To see that this is just the n-Yablo chain, consider the following renaming:

$$S_{(a,1)} \mapsto S_{(a-1)n+1},\ S_{(a,2)} \mapsto S_{(a-1)n+2},\ \ldots,\ S_{(a,n)} \mapsto S_{an}.$$

Then,

$$\begin{aligned} \delta^U(S_1) = \delta^U(S_{(1,1)}) &= \neg(S_{(1,2)}) \wedge \neg(S_{(2,2)}) \wedge \neg(S_{(3,2)}) \wedge \cdots \\ &= \neg(S_2) \wedge \neg(S_{2+n}) \wedge \neg(S_{2+2n}) \wedge \cdots \\ \delta^U(S_2) = \delta^U(S_{(1,2)}) &= \neg(S_{(1,3)}) \wedge \neg(S_{(2,3)}) \wedge \neg(S_{(3,3)}) \wedge \cdots \\ &= \neg(S_3) \wedge \neg(S_{3+n}) \wedge \neg(S_{3+2n}) \wedge \cdots \\ &\vdots \\ \delta^U(S_{n-1}) = \delta^U(S_{(1,n-1)}) &= \neg(S_{(1,n)}) \wedge \neg(S_{(2,n)}) \wedge \neg(S_{(3,n)}) \wedge \cdots \\ &= \neg(S_n) \wedge \neg(S_{2n}) \wedge \neg(S_{3n}) \wedge \cdots \\ \delta^U(S_n) = \delta^U(S_{(1,n)}) &= \neg(S_{(2,1)}) \wedge \neg(S_{(3,1)}) \wedge \neg(S_{(4,1)}) \wedge \cdots \\ &= \neg(S_{n+1}) \wedge \neg(S_{2n+1}) \wedge \neg(S_{3n+1}) \wedge \cdots \\ \delta^U(S_{n+1}) = \delta^U(S_{(2,1)}) &= \neg(S_{(2,2)}) \wedge \neg(S_{(3,2)}) \wedge \neg(S_{(4,2)}) \wedge \cdots \\ &= \neg(S_{n+2}) \wedge \neg(S_{2n+2}) \wedge \neg(S_{3n+2}) \wedge \cdots \\ &\vdots \end{aligned}$$

□

Corollary 6.15. *The n-cycle is paradoxical if and only if the n-Yablo chain is paradoxical.*

6.2. **Connecting paradox and contextuality.** At a conceptual level, semantic paradox and strong contextuality seem like the same kind of phenomenon—both are cases of the non-existence of global assignments. To formalize this connection, we demonstrate that one can associate empirical models to sets of sentences in such a way that the sentences are paradoxical if and only if the associated empirical model is strongly contextual.

Definition 6.16. Given a finite set of sentences $\{S_i\}_{i\leq n}$, define the associated event presheaf $\mathcal{E} : \mathcal{P}(X)^{\text{op}} \to \mathbf{Set}$ by setting $X := \{S_i\}_{i\leq n}$ and $O := T = \{0,1\}$. In this context, for a set of sentences $U \subseteq X$, $\mathcal{E}(U)$ is the set of possible (not necessarily acceptable) truth-value assignments $\sigma : U \to \{0,1\}$.

Definition 6.17. Given also a denotation function δ, consider the cover $\mathcal{M}$ of X given by

$$\mathcal{M} = \{C_i\}_{i \leq n}, \text{ where } C_i \coloneqq \{S_i\} \cup D_\delta(S_i).$$

Define the associated empirical model $\mathcal{S} \subseteq \mathcal{E}$ such that for all $C \in \mathcal{M}$,

$$\mathcal{S}(C) \coloneqq \{\sigma \in \mathcal{E}(C) \mid \sigma \text{ is acceptable}\}.$$

Proposition 6.18. *$\{S_i\}_{i \leq n}$ with δ is paradoxical if and only if the associated empirical model $\mathcal{S}$ is strongly contextual.*

Proof. This follows immediately from the fact that an element of $\mathcal{S}(X)$ is an acceptable truth-value assignment on $\{S_i\}_{i \leq n}$. ☐

Moreover, we can connect n-cycles and n-Yablo chains to the previous discussion of cohomological strong contextuality.

Proposition 6.19. *If n is odd, then the empirical model associated to the n-cycle is an all-versus-nothing model.*

Proof. Let $\{S_i\}_{i \leq n}$ with δ be the n-cycle with odd n. Recall (Definition 6.5) that δ is defined by

$$\delta(S_i) = \begin{cases} \neg(S_{i+1}) & i < n \\ \neg(S_1) & i = n \end{cases}.$$

Define $\mathcal{M}$ according to Definition 6.17. That is, $\mathcal{M} = \{C_i\}_{i \leq n}$ where for all $i < n$,

$$C_i = \{S_i, S_{i+1}\},$$

and for $i = n$,

$$C_n = \{S_n, S_1\}.$$

Then, $\mathcal{M}$ defines the associated empirical model $\mathcal{S}$. Recall (Definition 5.3) that an all-versus-nothing model depends on a ring R. We will use the boolean ring $\mathbb{B}$. Then, $\mathcal{S}$ determines a $\mathbb{B}$-linear theory

$$\mathbb{T}_\mathbb{B}(\mathcal{S}) = \bigcup_{i \leq n} \mathbb{T}_\mathbb{B}(\mathcal{S}(C_i)).$$

(See Definitions 5.1 and 5.2.) Note that for $i < n$,

$$\mathbb{T}_\mathbb{B}(\mathcal{S}(C_i)) = \{\langle C_i, (0,0), 0\rangle, \langle C_i, (1,1), 1\rangle\},$$

where a function $a : C_i \to \mathbb{B}$ is denoted by the ordered pair $(a(S_i), a(S_{i+1}))$. Similarly,

$$\mathbb{T}_\mathbb{B}(\mathcal{S}(C_n)) = \{\langle C_n, (0,0), 0\rangle, \langle C_n, (1,1), 1\rangle\},$$

where a function $a : C_n \to \mathbb{B}$ is denoted by $(a(S_n), a(S_1))$. Thus,

$$\mathbb{T}_\mathbb{B}(\mathcal{S}) = \bigcup_{i \leq n} \{\langle C_i, (0,0), 0\rangle, \langle C_i, (1,1), 1\rangle\}.$$

Suppose for sake of contradiction that $\mathcal{S}$ is not $\mathsf{AvN}_\mathbb{B}$, i.e. that there exists a map $g : \{S_i\}_{i \leq n} \to \mathbb{B}$ such that for all $\phi = \langle C_i, a, b\rangle \in \mathbb{T}_\mathbb{B}(\mathcal{S})$, $g|_{C_i}$ satisfies ϕ. Any g will satisfy equations of the form $\langle C_i, (0,0), 0\rangle$, so we only need to consider the equations of the form $\langle C_i, (1,1), 1\rangle$. For g to satisfy equations of that form for all $C_i \in \mathcal{M}$, g must satisfy the following system of equations:

$$\begin{array}{rcl} g(S_1) + g(S_2) & = & 1 \\ g(S_2) + g(S_3) & = & 1 \\ \vdots & = & \vdots \\ g(S_{n-1}) + g(S_n) & = & 1 \\ g(S_n) + g(S_1) & = & 1 \end{array}$$

Each term appears on the left side twice, so the left sides of the equations sum to 0. However, n is odd, so the right sides sum to 1. This is a contradiction, so no such g exists. Thus, $\mathsf{AvN}_{\mathbb{B}}(\mathcal{S})$. □

Corollary 6.20. *If n is odd, then the empirical model associated to the n-cycle is cohomologically strongly contextual.*

Denote that a set of sentences S with δ is paradoxical with $\mathsf{Pdox}(S, \delta)$ and that it forms an n-cycle with $\mathsf{Cycle}_n(S, \delta)$. Then, for odd n, we have

$$\mathsf{Cycle}_n(S, \delta) \Rightarrow \mathsf{CSC}_{\mathbb{B}}(\mathcal{S}) \Rightarrow \mathsf{SC}(\mathcal{S}) \Leftrightarrow \mathsf{Pdox}(S) \Leftrightarrow \mathsf{Pdox}(S^U)$$

where $\mathcal{S}$ is the empirical model associated with S and S^U is the unwinding of S, or the n-Yablo chain. So in the case of semantic paradoxes, cohomological witnesses of contextuality serve also as witnesses of paradox.

7. Digraphs, games, and more

Finally, we give a very short introduction into how contextuality may arise in the study of directed graphs. Cook [6] demonstrates that a set of sentences is paradoxical if a certain directed graph has no kernels. Taking this as our inspiration, we consider the connection between the non-existence of kernels in digraphs and strong contextuality.

Definition 7.1. A *directed graph* (digraph) $G = (V, E)$ consists of a set V of vertices and set E of ordered pairs of vertices called edges.

Definition 7.2. A *kernel* $K \subseteq V$ is

i) *stable*, meaning that for all $x, y \in K$, $(x, y) \notin E$;
ii) *absorbing*, meaning that for all $x \in V/K$, there is a $y \in K$ such that $(x, y) \in E$.

Definition 7.3. The set $N^+(x)$ of *out neighbors* of a vertex $x \in V$ is given by

$$N^+(x) \coloneqq \{y \in V \mid (x, y) \in E\}.$$

Definition 7.4. Given a graph G with $|V| = n$, enumerate the vertices $V = \{v_1, v_2, \ldots, v_n\}$ and define the associated event presheaf by $X \coloneqq V$ and $O = \{0, 1\}$. Define a measurement cover by

$$\mathcal{M} = \{C_i\}_{i \leq n} \text{ where } C_i \coloneqq \{v_i\} \cup N^+(v_i).$$

Define the associated empirical model $\mathcal{S}$ by letting

$$\mathcal{S}(C) \coloneqq \{\sigma \in O^C \mid \text{for all } x \in C,\ \sigma(x) = 1 \text{ iff for all } y \in C \cap N^+(x),\ \sigma(y) = 0\}.$$

Proposition 7.5. *G contains no kernels if and only if $\mathcal{S}$ is strongly contextual.*

Proof. Suppose that $\mathcal{S}$ is strongly contextual and suppose for sake of contradiction that G contains a kernel K. Define a map $g : X \to O$ by

$$g(x) = \begin{cases} 0 & x \notin K \\ 1 & x \in K \end{cases}.$$

We need to check that $g \in \mathcal{S}(X)$. Suppose $g(x) = 1$. Then, $x \in K$, so $N^+(x) \subseteq X/K$. Thus, for all $y \in N^+(x)$, $g(y) = 0$. Alternatively, fix x and suppose that all $y \in N^+(x)$ satisfy $g(y) = 0$. Then, $N^+(x) \subseteq X/K$. If $x \in X/K$, then there is at least one $z \in N^+(x)$ such that $z \in K$, so x must be in K, so $g(x) = 1$. Thus, $g \in \mathcal{S}(X)$, so $\mathcal{S}$ is not strongly contextual. By contradiction, we have that G contains no kernels.

Suppose instead that G contains no kernels and suppose for sake of contradiction there exists a global section $g \in \mathcal{S}(X)$. Consider the set $g^{-1}(1)$. Note that if $x, y \in g^{-1}(1)$, then $(x, y) \notin E$. (Else, $y \in N^+(x)$, so $g(y) = 0$, so $y \notin g^{-1}(1)$.) Then, consider $x \notin g^{-1}(1)$. There must be some $y \in g^{-1}(1)$ such that $(x, y) \in E$. (Else, $N^+(x) \subseteq g^{-1}(0)$, which means that $g(x) = 1$.) Thus, $g^{-1}(1)$ is a kernel. By contradiction, we have that no $g \in \mathcal{S}(X)$ exists, so $\mathcal{S}$ is strongly contextual. □

The existence of a kernel in a directed graph has to do with the existence of a solution to a two-person game on the graph [14] and can be used to study a variety of situations where players alternate turns until one player is out of moves. Thus, we can speculate on some degree of connection between strongly contextual empirical models, semantic paradoxes, and unsolvable games.

Contextuality forces us to confront the strangeness of quantum mechanics in non-intuitive ways. Perhaps by understanding how contextuality arises in our macroscopic world as well as in the microscopic one, we can come to better understand the mathematical theory and physical reality that lies beneath everything around us. For example, there may very well be usefulness in an analogy between games and contextuality, and this idea has been studied by others [1]. Or perhaps the fact that quantum reality verges on paradox suggests a new perspective on logic—the progenitors of topos quantum theory, for instance, prefer intuitionistic logic.

The careful analysis of all the facts and consequences of quantum contextuality is not completed, and there is more to discover about this part of the mathematical and physical worlds. However, much interesting progress has made so far, and we hope that this paper is an enlightening survey of where the field currently stands and how we got here.

Acknowledgments

Thanks to my mentor Alicia Lima for her helpful expertise and support throughout the program. Thanks as well to Peter May and everyone who worked to make the REU such a great experience.

References

[1] Samson Abramsky, Rui Soares Barbosa, and Amy Searle. Combining contextuality and causality: a game semantics approach. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 382(2268), January 2024.

[2] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015.

[3] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964.
[4] David Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Phys. Rev.*, 85:166–179, Jan 1952.
[5] Adán Cabello, JoséM. Estebaranz, and Guillermo García-Alcaine. Bell-kochen-specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, March 1996.
[6] Roy T. Cook. Patterns of paradox. *The Journal of Symbolic Logic*, 69(3):767–774, 2004.
[7] A. Döring and C. Isham. *"What is a Thing?": Topos Theory in the Foundations of Physics*, page 753–937. Springer Berlin Heidelberg, 2010.
[8] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10):777, 1935.
[9] Cecilia Flori. *A First Course in Topos Quantum Theory*. Springer Berlin Heidelberg, 2013.
[10] LUCIEN HARDY. Are quantum states real? *International Journal of Modern Physics B*, 27(01n03):1345012, November 2012.
[11] Simon Kochen and Ernst P Specker. The problem of hidden variables in quantum mechanics. *Ernst Specker Selecta*, pages 235–263, 1990.
[12] N. David Mermin. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.*, 65:3373–3376, Dec 1990.
[13] A. Pais. Einstein and the quantum theory. *Rev. Mod. Phys.*, 51:863–914, Oct 1979.
[14] Alvin E Roth. Two-person games on graphs. *Journal of Combinatorial Theory, Series B*, 24(2):238–241, 1978.
[15] Stephen Yablo. Paradox without self-reference. *Analysis*, 53(4):251–252, 1993.