BELL'S THEOREM IN CATEGORICAL QUANTUM MECHANICS

AJ DEROSA

Abstract. We demonstrate the viability of using category theory to explore the foundations of quantum theory. To do so, we use category theory to prove Bell's theorem, which asserts that quantum mechanics does not admit a refinement as a local hidden variables theory. Assuming no knowledge of category theory or quantum theory, we work from basic categorical notions to construct a category of quantum processes. Using this construction, we prove that quantum processes are non-signaling. Finally, we use a thought experiment to give a non-probabilistic proof of Bell's theorem and discuss the possible mathematical origins of non-locality.

Contents

1. Introduction

Applied category theory is an emerging field that uses the abstract tools of category theory to model real-world phenomena. Categorical quantum mechanics—a sub-field of applied category theory—has yielded substantial results in the theories of quantum computing and quantum foundations.

We demonstrate the promise of using categorical quantum mechanics to investigate the foundations of quantum theory by proving and analyzing Bell's theorem. Bell's theorem places a restriction on using hidden variables theories to refine quantum mechanics. A hidden variables theory is a theory that replicates the results of quantum mechanics and is *realistic*, which means that the results of measurements are stored by variables that exist regardless of whether or not measurements occur. As we will see, quantum mechanics is not a realistic theory, but it can be refined into a hidden variables theory that is realistic. Bell's theorem shows that such a hidden variables theory must be *non-local*, which means that it allows for the instantaneous transfer of information—a contradiction of special relativity.

In Section 2, we construct a category of quantum processes by defining †-hypergraph categories and then specifying some "quantum" structure. In Section 3,

Date: August 29, 2023.

we discuss causal structures and prove that quantum processes are non-signaling; then, we consider parity calculations for a GHZ-Mermin thought experiment to derive a contradiction between categorical quantum mechanics and local hidden variables theories. In Section 4, we conclude with a brief discussion of the meaning of Bell's theorem and avenues of further exploration using categorical quantum mechanics.

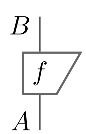
2. Categorical quantum mechanics

Broadly, our aim is to construct a theory of processes. The most basic thing we should be able to model with such a theory is performing one process after another. That is, we want to *compose* processes. An emphasis on composition is inherent in category theory, so we begin there.

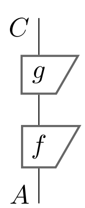
Definition 2.1. A *category* $\mathcal{C}$ consists of a collection of objects $\mathrm{ob}(\mathcal{C}) = \{A, B, C, ...\}$ and a collection of morphisms $f, g, h, ...$ A morphism $f : A \to B$ maps object A to object B. And for every object A, there is an identity morphism $1_A : A \to A$. For morphisms $f : A \to B$, $g : B \to C$, there exists the composite morphism $gf : A \to C$. (Read "g following f.") Composition is subject to two axioms:

(1) *Unitality*: For any morphism $f : A \to B$,
$$1_B f = f = f 1_A.$$
(2) *Associativity*: For any morphisms $f : A \to B$, $g : B \to C$, $h : C \to D$,
$$(hg)f = h(gf) =: hgf.$$

We denote categorical notions with *wire diagrams*. Morphisms are denoted as boxes, and objects are denoted as wires going into and out of boxes. In this paper, we read diagrams from bottom to top, so $f : A \to B$ is denoted as follows.

(2.2)  $:= \quad f : A \to B$

Given $f : A \to B$ and $g : B \to C$, we denote the composite morphism $gf : A \to C$ by connecting boxes as follows. And an identity morphism is just a plain wire.

(2.3)  $:= \quad gf : A \to C$  $:= \quad 1_A : A \to A$

Another basic thing we should be able to model is performing two processes "at the same time." For this, we need a particular type of category, for which we need the following two preliminary notions.

Definition 2.4. A *functor* $F : \mathcal{C} \to \mathcal{D}$ is a map between categories $\mathcal{C}, \mathcal{D}$. A functor consists of an object $F(C) \in \mathrm{ob}(\mathcal{D})$ for every object $C \in \mathrm{ob}(\mathcal{C})$ and a morphism $F(f) : F(C) \to F(C')$ in $\mathcal{D}$ for every morphism $f : C \to C'$ in $\mathcal{C}$. Functors are subject to two axioms:

(1) *Distributivity over composition*: For all f, g in $\mathcal{C}$ such that there exists gf in $\mathcal{C}$, we have $F(g)F(f) = F(gf)$.
(2) *Identities map to identities*: For all $C \in \mathrm{ob}(\mathcal{C})$, we have $F(1_C) = 1_{F(C)}$.

Definition 2.5. Given categories $\mathcal{C}, \mathcal{D}$, there exists a *product category* $\mathcal{C} \times \mathcal{D}$. For every pair of objects $C \in \mathrm{ob}(\mathcal{C})$, $D \in \mathrm{ob}(\mathcal{D})$, there is an object $(C, D) \in \mathrm{ob}(\mathcal{C} \times \mathcal{D})$. For every pair of morphisms $f : C \to C'$ in $\mathcal{C}$, $g : D \to D'$ in $\mathcal{D}$, there is a morphism $(f, g) : (C, D) \to (C', D')$ in $\mathcal{C} \times \mathcal{D}$.

Now we can define the category we need.

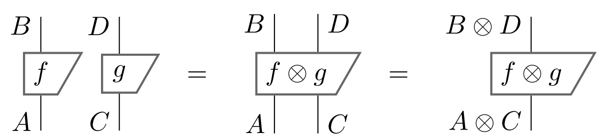
Definition 2.6. A *symmetric monoidal category* $(\mathcal{C}, I, \otimes)$ consists of a category $\mathcal{C}$, unit object I, and functor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ subject to the following conditions. (Note we use infix notation to refer to the functor, so $\otimes$ maps the object (A, B) to the object denoted $A \otimes B$.)

(1) *Unitality*: For all $A \in \mathrm{ob}(\mathcal{C})$, there exists an isomorphism $\lambda_A : I \otimes A \cong A$. Similarly, there exists an isomorphism $\rho_A : A \otimes I \cong A$.
(2) *Associativity*: For all $A, B, C \in \mathrm{ob}(\mathcal{C})$, there exists an isomorphism $\alpha_{A,B,C}$: $(A \otimes B) \otimes C \cong A \otimes (B \otimes C)$.
(3) *Symmetry*: For every $A, B \in \mathrm{ob}(\mathcal{C})$, there exists an isomorphism $\sigma_{A,B}$: $A \otimes B \cong B \otimes A$.

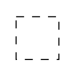
In the case of a *strict* symmetric monoidal category, the isomorphisms are equalities.

And similarly for morphisms. To denote morphisms $f \otimes g$, we draw the boxes in parallel or as one composite box. Similarly, objects $A \otimes C$ are denoted as two parallel wires or one composite wire.

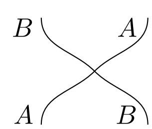
(2.7)



The unit object is the empty diagram.



Finally, symmetry lets us swap objects, which we denote as swapping wires.



This is a fairly vanilla category of processes. To make things quantum, we'll need a way to consider *entanglement*, for which we'll need non-$\otimes$-separable morphisms. These come with a *compact closed category*.

Definition 2.8. Given symmetric monoidal category $(\mathcal{C}, I, \otimes)$ and object $A \in \mathrm{ob}(\mathcal{C})$, a *dual* of A consists of an object $A^* \in \mathrm{ob}(\mathcal{C})$, a *unit* morphism $\eta_A : I \to A^* \otimes A$, and a *counit* morphism $\epsilon_A : A \otimes A^* \to I$ such that the following diagrams commute.

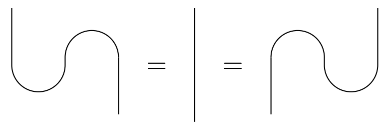
$$
\begin{tikzcd}
A \otimes I \arrow[d, "1_A \otimes \eta_A"'] \arrow[dr, "1_A"] & \\
A \otimes A^* \otimes A \arrow[r, "\epsilon_A \otimes 1_A"'] & I \otimes A
\end{tikzcd}
\qquad
\begin{tikzcd}
I \otimes A^* \arrow[d, "\eta_A \otimes 1_{A^*}"'] \arrow[dr, "1_{A^*}"] & \\
A^* \otimes A \otimes A^* \arrow[r, "1_{A^*} \otimes \epsilon_A"'] & A^* \otimes I
\end{tikzcd}
\tag{2.9}
$$

$\mathcal{C}$ is a *compact closed category* if it contains a dual object A^* for every object $A \in \mathcal{C}$. Objects are *self-dual* if $A^* = A$.

We denote the unit and counit as a *cup* and *cap*, respectively.

Thus, (2.9) is equivalent to the following identity.

(2.10)



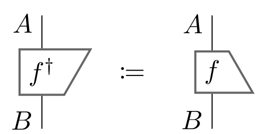
The next basic thing we ask from our theory is that we can "undo" processes. For this, we need a *dagger*, for which we require a preliminary definition.

Definition 2.11. Given a category $\mathcal{C}$, there exists an *opposite category* $\mathcal{C}^{\mathrm{op}}$, which satisfies $\mathrm{ob}(\mathcal{C}^{\mathrm{op}}) = \mathrm{ob}(\mathcal{C})$, and for all morphisms $f : A \to B$ in $\mathcal{C}$, there exists a morphism $f^{\mathrm{op}} : B \to A$ in $\mathcal{C}^{\mathrm{op}}$.

Definition 2.12. A $\dagger$-*category* (read as "dagger category") $(\mathcal{C}, \dagger)$ consists of a category $\mathcal{C}$ and a functor $\dagger : \mathcal{C} \to \mathcal{C}^{\mathrm{op}}$ such that $\dagger$ maps a morphism $f : A \to B$ in $\mathcal{C}$ to $f^\dagger : B \to A$ in $\mathcal{C}^{\mathrm{op}}$. $\dagger$ is subject to two axioms.

(1) *Identity on objects*: $A^\dagger = A$.
(2) *Involution on morphisms*: $(f^\dagger)^\dagger = f$.

Given a morphism $f : A \to B$, we denote $f^\dagger$ as a vertically reflected box. (Additionally, we refer to $f^\dagger$ as the *adjoint* of f.)



Remark 2.13. $f^\dagger : B \to A$ does not necessarily "undo" $f : A \to B$ in the sense that it is not necessarily the case that $f^\dagger f = 1_A$. When this is the case, we say that f is an *isometry*. If it is also the case that $f f^\dagger = 1_B$, we say that f is *unitary*.

Definition 2.14. A $\dagger$-*compact closed category* $(\mathcal{C}, I, \otimes, \dagger)$ consists of a compact closed category $(\mathcal{C}, I, \otimes)$ and a functor $\dagger$ that satisfy the following compatibility conditions.

(1) *Distributivity over monoidal functor*: $(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$.
(2) *Dagger of counit is dual unit*: $\epsilon_A^\dagger = \eta_{A^*}$.

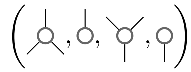
We now add some structure to our objects, which essentially ensures that "copying" and "deleting" data follow intuitive rules.

Definition 2.15. A $\dagger$-*special commutative Frobenius algebra* ($\dagger$-SCFA) $(A, \mu, \eta, \delta, \epsilon)$ consists of an object A, multiplication $\mu : A \otimes A \to A$, unit $\eta : I \to A$, comultiplication $\delta : A \to A \otimes A$, and counit $\epsilon : A \to I$ that satisfy the following axioms.

(1) *Associativity*: $\mu(\mu \otimes 1) = \mu(1 \otimes \mu)$.
(2) *Unitality*: $\mu(\eta \otimes 1) = 1 = \mu(1 \otimes \eta)$.
(3) *Coassociativity*: $(\delta \otimes 1)\delta = (1 \otimes \delta)\delta$.
(4) *Counitality*: $(\epsilon \otimes 1)\delta = 1 = (1 \otimes \epsilon)\delta$.
(5) *Frobenius law*: $(1 \otimes \mu)(\delta \otimes 1) = \delta\mu = (\mu \otimes 1)(1 \otimes \delta)$.

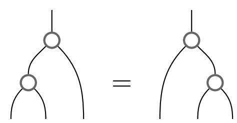
(6) *Commutativity*: compatible with symmetry.
(7) *Special law*: $\mu\delta = 1$.

In our wire diagrams, we denote these morphisms as follows.

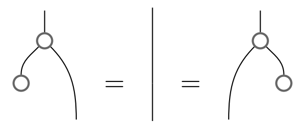
$(\mu, \eta, \delta, \epsilon) =$ 

In diagram form, the axioms of a †-SCFA are as follows.

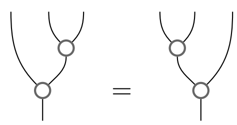
(1)



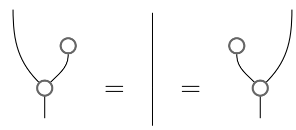
(2)



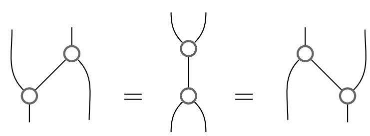
(3)



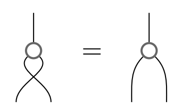
(4)



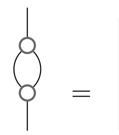
(5)



(6)

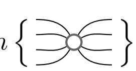


(7)



We can now replace wire diagrams with the far-more-interesting *spider diagrams.*

Definition 2.16. Given a Frobenius algebra $(A, \mu, \eta, \delta, \epsilon)$ where A is in a symmetric monoidal category $(\mathcal{C}, I, \otimes)$, a *spider* $\bigcirc_m^n : A^{\otimes m} \to A^{\otimes n}$ is defined to be the morphism given by applying μ composed with itself $(m-1)$ times followed by applying δ composed with itself $(n-1)$ times. We denote this as a node with m input wires and n output wires.

m  n

We use this notation because it does not actually matter how we compose various instances of $\mu, \eta, \delta, \epsilon$ to create a morphism. All that matters is the number of inputs and outputs. We formalize this with the following theorem.

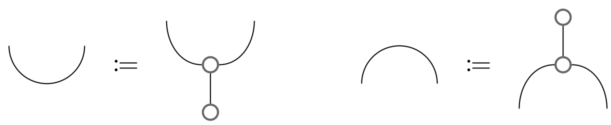
Theorem 2.17. *Fix a Frobenius algebra $(A, \mu, \eta, \delta, \epsilon)$ in a symmetric monoidal category $(\mathcal{C}, I, \otimes)$. Suppose we have the morphism $f : A^{\otimes m} \to A^{\otimes n}$ constructed only from $\mu, \eta, \delta, \epsilon$ using composition and the monoidal functor $\otimes$ and containing only one connected component. Then, $f = \bigcirc_m^n$.*

Replacing all wires with spiders, we end up with a useful type of category.

Definition 2.18. A †-*hypergraph category* consists of a †-symmetric monoidal category $(\mathcal{C}, I, \otimes, \dagger)$ such that every object $A \in \mathrm{ob}(\mathcal{C})$ is equipped with a Frobenius algebra along with certain compatibility conditions.

Recall the cups and caps from Definition 2.8. We define them for a hypergraph category as follows.

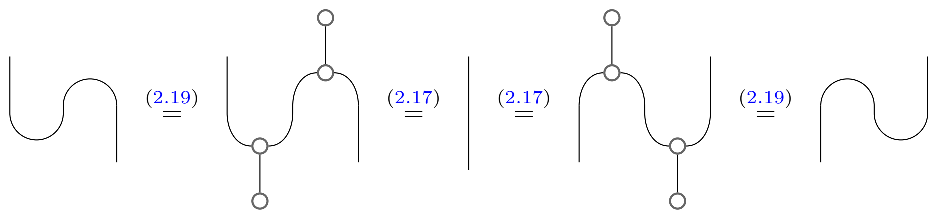
(2.19)



We just need to check that they satisfy (2.10).

Theorem 2.20. *A hypergraph category with self-dual objects and units and counits defined as in 2.19 is compact closed.*

Proof.

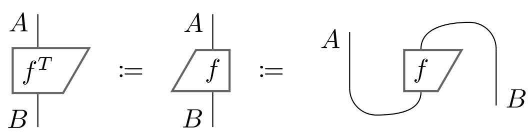


□

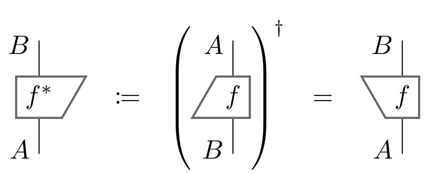
But what use is this sort of category for quantum theory? Quantum processes are traditionally represented as matrices, and what do matrices have to do with †-hypergraph categories? It turns out that $\mathbf{Mat}(\mathbb{C})$—the category whose objects are natural numbers and whose morphisms $M : m \to n$ are $n \times m$ matrices—*is* a †-hypergraph category[1].

With the next two definitions, we generalize notions that typically pertain to matrices.

Definition 2.21. We define a *transpose* functor $T : \mathcal{C} \to \mathcal{C}$. T is the identity on objects. And given a morphism $f : A \to B$, the transpose $f^T : B \to A$ is defined as follows, and its box is denoted as a 180° rotation of f.



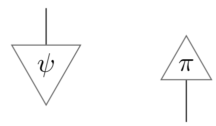
Definition 2.22. We define a *conjugate* functor $* : \mathcal{C} \to \mathcal{C}$. $*$ is the composite functor $\dagger T$. Thus, it is the identity on objects and maps morphisms to their adjoint transpose. We denote the box for conjugate morphisms as a horizontal reflection of f.



[1]We will not get into the details of this fact here, but we refer the interested reader to [2].

And we can generalize a few more notions from the standard formalism of quantum mechanics.

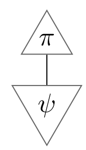
Definition 2.23. A *state* $\psi : I \to A$ is a morphism from the unit object. Similarly, an *effect* $\pi : A \to I$ is a morphism to the unit object. A *scalar* $\lambda : I \to I$ is a morphism from the unit object to itself. We have the following notations.



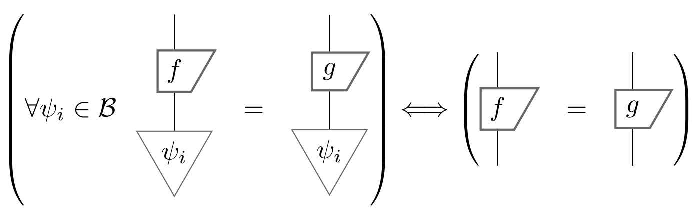
$$\lambda$$

Remark 2.24. The *Born rule* of quantum mechanics typically states that the probability of an effect π given a state ψ is given by the inner product $\langle \pi \mid \psi \rangle$. We have a different way of getting a scalar from a state and effect: composition.[2]

$$P(\pi \mid \psi) \quad =$$



Definition 2.25. A *basis* $\mathcal{B} = \{\psi_i : I \to A\}_i$ is a set of states such that for any morphisms $f, g : A \to B$, if for all $\psi_i \in \mathcal{B}$ we have that $f\psi_i = g\psi_i$, then $f = g$ and vice versa. That is,



We say a basis with $|\mathcal{B}| = D$ has *dimension* D.

Definition 2.26. Given objects A, B, a *zero morphism* is a morphism $0_{AB} : A \to B$ such that for any morphism $f : A \to B$,

$$0_{BB} f = f 0_{AA} = 0_{AB} \qquad\qquad 0_{CD} \otimes f = 0_{(C\otimes A)(D\otimes B)}$$

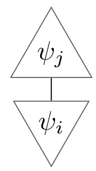
We use zero morphisms to model impossible processes. Intuitively, if any constituent of a process is not possible, then the whole process is not possible. This is consistent with the above definition: if a diagram contains a single zero morphism, the entire diagram is a zero morphism. We may omit the subscript on zero morphisms, understanding the (co)domain from context. Thinking of zero morphisms as the analogue of the number 0 and the unit morphism as the analogue of the number 1, we have an intuitive definition of orthonormal bases.

Definition 2.27. A basis is *orthonormal* if for all $\psi_i, \psi_j \in \mathcal{B}$,

$$\psi_j^\dagger \psi_i = \begin{cases} I & \text{if } i = j \\ 0 & \text{else} \end{cases} .$$

[2]This yields a probability for quantum processes (which we haven't defined yet) but not in general.

That is,



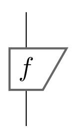
$$= \begin{cases} I & \text{if } i = j \\ 0 & \text{else} \end{cases}.$$

Definition 2.28. A category *enriched in commutative monoids* consists of a category $\mathcal{C}$ and for every pair of objects A, B a zero morphism 0_{AB} and a commutative monoid $(\mathcal{C}(A,B), +, 0)$[3] that is compatible with composition. That is, for morphisms $f, g, h : A \to B$, $j : B \to C$, and $k : D \to A$ the following properties are satisfied.

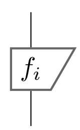
(1) *Associativity*: $(f+g)+h = f+(g+h)$
(2) *Commutativity*: $f+g = g+f$
(3) *Unitality*: $f+0 = f = 0+f$
(4) *Zero property*: $0f = 0 = g0$
(5) *Left distributivity*: $j(f+g) = (jf)+(jg)$
(6) *Right distributivity*: $(g+f)k = (gk)+(hk)$

This structure enables us to take sums of morphisms. We can then define probabilistic processes as sums of morphisms.

(2.29)



$$= \sum_{i=1}^{n} p_i$$

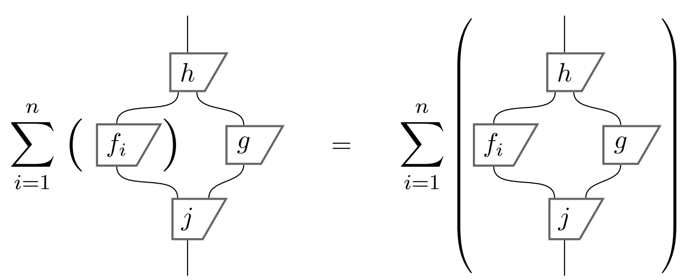


$$:= \quad (p_1 \otimes f_1) + (p_2 \otimes f_2) + \cdots + (p_n \otimes f_n)$$

Remark 2.30. As is unavoidable in quantum mechanics, we will deal with probabilities throughout this paper. However, our main proof will not be probabilistic, so we are content to omit the scalars p_i and prove things *up to a scalar*. To denote equality up to a scalar, we use $\equiv$.

From the distributivity properties in Definition 2.28, we deduce that sums distribute over a diagram. That is, if a diagram contains at least one probabilistic process, then the entire diagram is probabilistic.

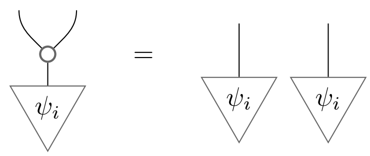
Example 2.31.



It's certainly a nice bonus that this structure gives us a notion of probabilistic processes—especially since they are a central part of quantum mechanics. But we introduced the idea of a zero process to define orthonormal bases (ONBs). Recall that in a †-hypergraph category, every object is equipped with a †-SCFA; it turns out that there is a sort of equivalence between ONBs and †-SCFAs.

[3]The notation $\mathcal{C}(A,B)$ denotes the set of morphisms $A \to B$.

Theorem 2.32. *For any †-SCFA, the set of states* $\{\psi_i\}_i$ *such that*

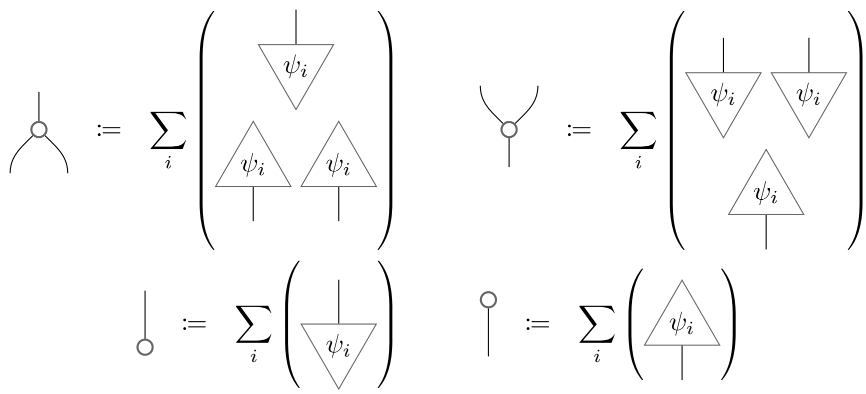


is an ONB; every †-SCFA uniquely determines and is uniquely determined by an ONB.

A sketch of the proof is given in [2].

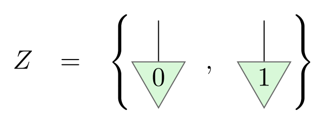
This means we can define the constituents of an †-SCFA in terms of an ONB. Suppose we have an ONB $\mathcal{B} = \{\psi_i\}_i$. Then, we define the associated †-SCFA as follows.

(2.33)



This relationship between †-SCFAs and ONBs is key in categorical quantum mechanics. We will use it to introduce the following quantum notions: phase, complementarity, and measurement. First, however, we make things more concrete, fixing dimension $D = 2$ and defining some bases. (Note we refer to two-dimensional quantum systems as *qubits*.) We have the Z basis, denoted in green, for qubits.

(2.34)

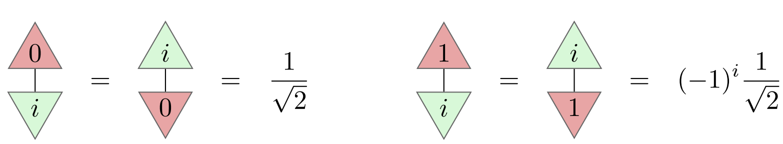


In terms of Z, we define the X basis, denoted in red.

(2.35)

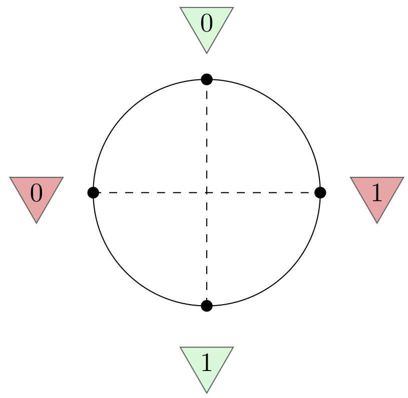
Thus, we have the following for Z basis state i.

(2.36)



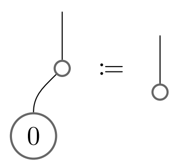
Definition 2.37. The *Bloch sphere*—a cross-section of which is shown below—is a geometric representation of the possible states of a qubit, in which each ONB is

represented as an axis.



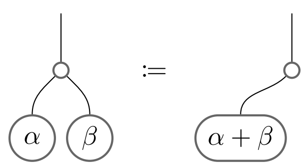
Definition 2.38. A *phase* $\alpha : I \to A$ is a special kind of state. One way to think about a phase is as a rotation about an axis of the Bloch sphere. Thus, the phases about any particular axis are isomorphic to the circle group $U(1)$. We define *phase spiders* (spiders that induce a phase on objects) with this group structure in mind. The group identity is a phase of 0.

(2.39)



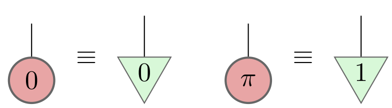
The group sum is defined as follows (mod 2π).

(2.40)



Remark 2.41. Using the Bloch sphere, we see that a phase of π about the X-axis maps the 0 state of the Z basis to the 1 state of the Z basis. Thus, we have the following facts[4].

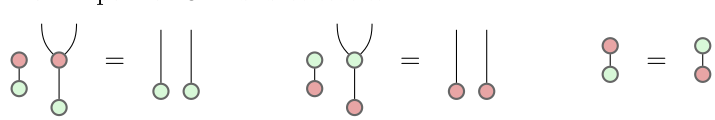
(2.42)



The Z and X bases are an example of a pair of bases with a special relationship. We classify pairs of ONBs as *coherent* and/or *complementary*. (The following definitions are general; red and green do not necessarily refer the Z and X bases.)

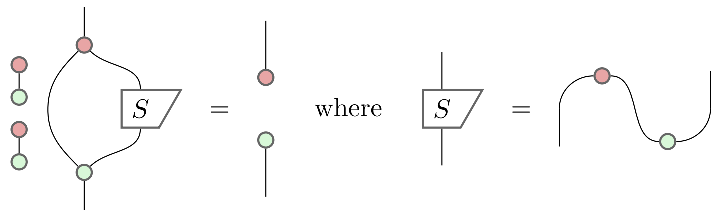
Definition 2.43. A pair of ONBs is *coherent* if

(2.44)



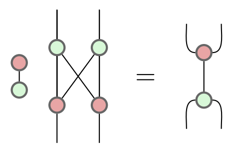
[4]We elide a full treatment of where these facts come from and again refer the interested reader to [2].

Definition 2.45. A pair of ONBs is *complementary* if



Definition 2.46. A pair of ONBs is *strongly complementary* if it is coherent and

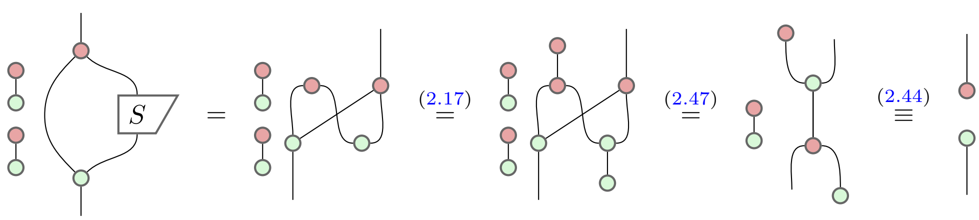
(2.47)



Example 2.48. The Z and X bases are strongly complementary. A proof follows directly from Definition 2.46 and (2.35).

Theorem 2.49. *Strongly complementary ONBs are complementary.*

Proof.

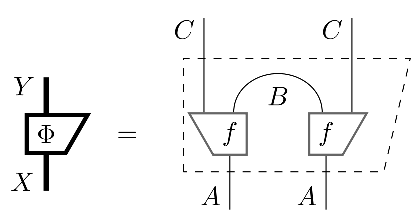


□

Definitions 2.43, 2.45, and 2.46 along with Theorem 2.49 indicate that strongly complementary ONBs form a scaled *Hopf algebra*. As discussed in Remark 2.30, we can omit the scalars and just work with the Hopf algebra equations.

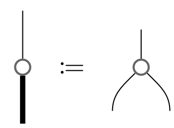
Finally, we distinguish between quantum and classical processes.

Definition 2.50. A *quantum process* $\Phi : X \to Y$ is a *completely positive map*. That is, for some objects A, B, C such that $X = A \otimes A$, $Y = C \otimes C$ and some morphism $f : A \to B \otimes C$, we have the following.



Accordingly, we represent all things quantum with doubled (bold) lines and all things classical with single lines.

Definition 2.51. A *measurement* is a map from a quantum wire to a classical wire. Measurement in some ONB is defined as μ in the associated †-SCFA.



Quantum processes satisfy an additional property: if we "discard" the outputs of a process, it is the same as discarding the inputs. We call such a property *causality*. We can formalize this property. Recall that every object A comes equipped with a counit $\epsilon : A \to I$. Thus, I is termed a *terminal object*. We model a discard process as a morphism to a terminal object. Then, we can formulate causality as follows.

Property 2.52. Let $!_A : A \to I$ denote the unique morphism from an object A to terminal object I. A morphism $f : A \to B$ is *causal* if $!_B f = !_A$

At last, we bring everything we have discussed together:

Definition 2.53. We specify the category of quantum processes $\textsc{qm}$ as follows.

(1) $\textsc{qm}$ is a †-hypergraph c.
  (a) $(\textsc{qm}, I, \otimes)$ is a symmetric monoidal category equipped with a dagger functor $\dagger : \textsc{qm} \to \textsc{qm}^{\mathrm{op}}$. (See Definitions 2.6 and 2.12.)
  (b) Every object is equipped with a †-SCFA. (See Definition 2.15.)
(2) Objects in $\textsc{qm}$ are self-dual; cups and caps are defined as in (2.19). By Theorem 2.20, $\textsc{qm}$ is compact closed.
(3) The set of morphisms consists of completely positive maps between objects (including morphisms comprising †-SCFAs). (See Definition 2.50.)
(4) $\textsc{qm}$ is enriched in commutative monoids. (See Definition 2.28.)
(5) For any †-SCFA, there is an associated phase group. (See Definition 2.38.)
(6) All morphisms satisfy Property 2.52.

3. Bell's theorem

In this section, we prove the following statement of Bell's theorem.

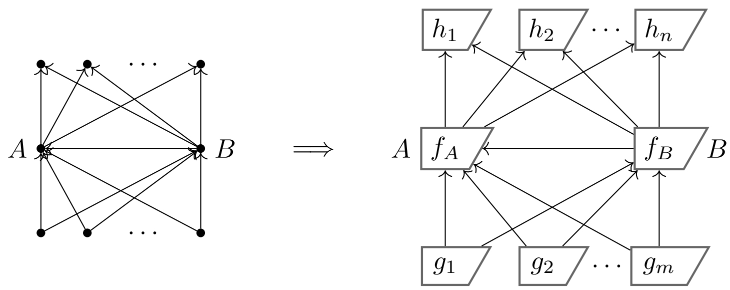
Theorem 3.1. *If a hidden variables theory is local, then it is inconsistent with quantum mechanics.*

Let's break down what this means. A hidden variable theory is a *realistic* refinement of quantum mechanics. By "realistic," we mean that the results of measurements are stored by variables that exist whether or not a measurement is performed. *Locality* refers to the consequence of special relativity that there is a maximum speed of information, so for some events A, B, there is no causal relationship $A \to B$.

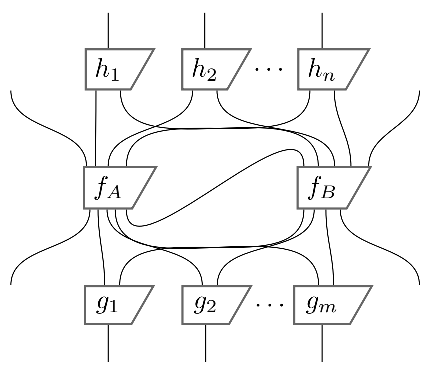
If quantum mechanics enabled communication between such events A and B (a phenomenon known as *signaling*), then Bell's theorem does not really give us any new or surprising information. So first, we prove that quantum mechanics is *non-signaling*. To do so, we represent a causal structure as a directed graph, where vertices are events and an arrow $A \to B$ indicates that A can affect B. We take the transitive closure of the graph because if A affects B and B affects C, we have that A affects C. Moreover, it must be that if $A \to B$, then there is no arrow $B \to A$. Thus, a causal structure is a transitive directed acylic graph.

Theorem 3.2. *If a symmetric monoidal category $\mathcal{C}$ has a terminal object T such that it satisfies Property 2.52, then it is non-signaling.*

Proof. We fix an arbitrary causal structure (arbitrary transitive directed acyclic graph). We then fix vertices A, B such that there does not exist an arrow $A \to B$. A and B have an arbitrary shared history, vertices C such that there exist arrows $C \to A$ and $C \to B$; an arbitrary shared future, vertices D such that there exist arrows $A \to D$ and $B \to D$; and possibly an arrow $B \to A$. We take the induced subgraph containing A and B as well as their shared history and shared future. We place arbitrary[5] morphisms in $\mathcal{C}$ at the vertices of this subgraph.



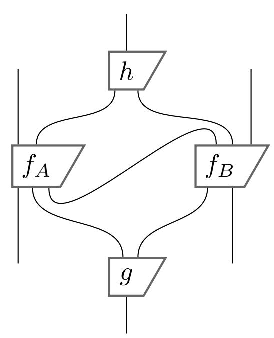
A causal arrow between morphisms means we can connect outputs to inputs to form the following wire diagram.



$= \quad (*)$

Let $h = h_1 \otimes h_2 \otimes \cdots \otimes h_n$, and let $g = g_1 \otimes g_2 \otimes \cdots \otimes g_m$. Then,

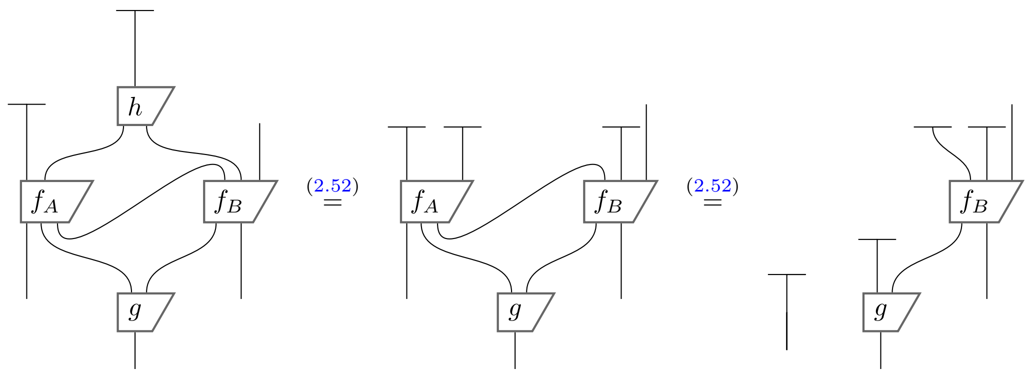
$(*) \quad =$



We now calculate what the output of f_B should be. We discard the outputs of h and f_A since they are not connected to f_B. Then, we apply Property 2.52 twice.

[5]Not entirely arbitrary—we require that they allow for the composition that comes next.

(Here, we denote a discard process $!_A : A \to T$ as a horizontal line.)



Thus, the output of f_B is not a function of the existence of f_A, so we conclude that $\mathcal{C}$ is non-signaling. □

Corollary 3.3. QM *is non-signaling.*

With that fact established, we have an interest in proving Bell's theorem. After all, if quantum mechanics was signaling and allowed information to be transmitted faster than the speed of light, it would seem obvious that any theory seeking to reproduce the results of quantum mechanics would require similar faster-than-light travel. Since we have just shown that quantum mechanics is non-signaling, we have no reason to expect that a hidden variables theory[6] must be non-local.

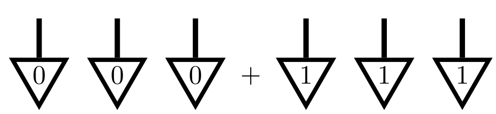
However, Bell's theorem tells us that this is not the case: indeed, any hidden variables theory that can replicate the predictions of quantum mechanics must be non-local. This is a substantial restriction to place on hidden variable theories, which are among many candidates for the correct way to formulate quantum mechanics.

Bell's original proof of this theorem involved probabilities; the advantage of his work and related probabilistic proofs is that they give us numerical values that can be (and have been) checked by experiments. However, these probabilistic proofs do not give much insight regarding what specific mathematical properties of quantum mechanics lead to such non-locality.

We present a non-probabilistic proof of Bell's theorem. Although it does not make for a great experimental design, it makes some headway on this fundamental question about non-locality.

We use a thought experiment involving *GHZ states.*

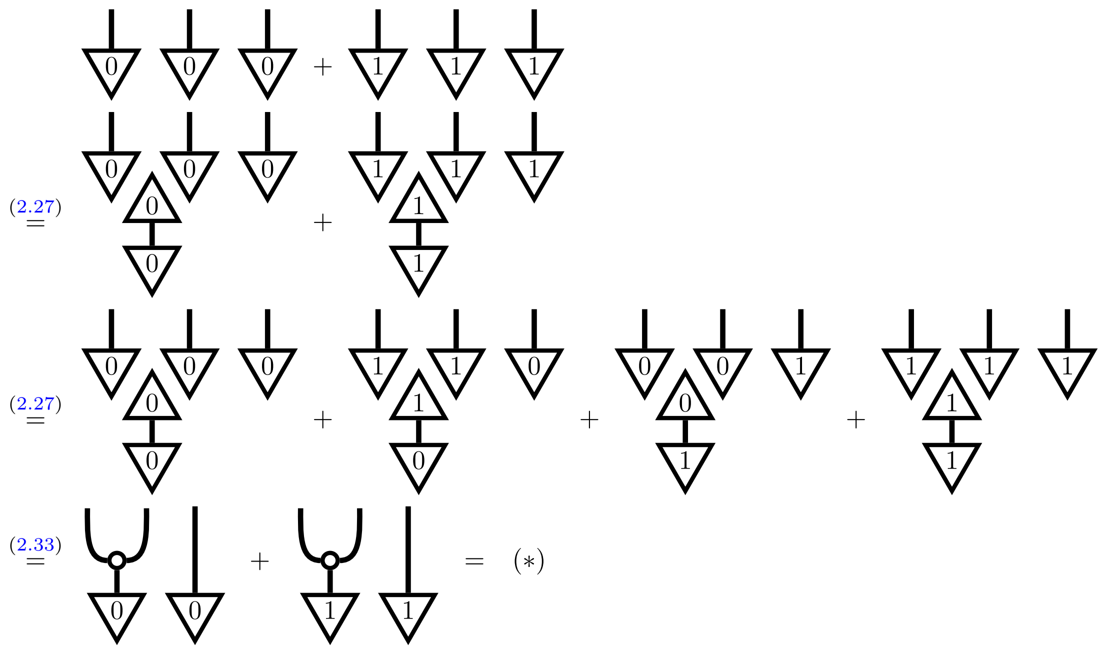
Definition 3.4. The *Greenberger–Horne–Zeilinger state* (GHZ state) is the three-qubit maximally entangled state, defined (up to a scalar) as follows.



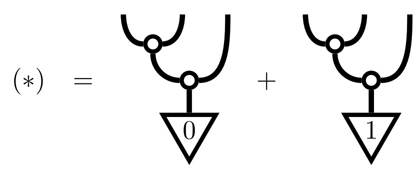
Lemma 3.5. *The GHZ state is equivalent to the spider* $\bigcirc_0^3$.

[6]a theory that replicates the results of quantum mechanics using realistic variables—variables that have a value regardless of if they are measured

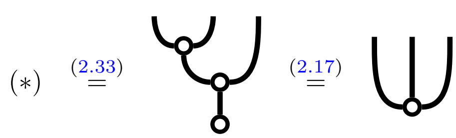
Proof. Recall from Definition 2.27 that composing states in an ONB yields either 0 or I. Then, we recall (2.33) to substitute in spiders.



We repeat the same process one more time, which gives us

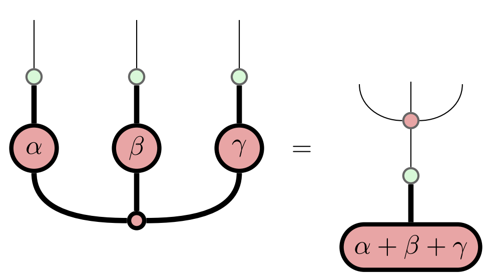


We use (2.33) to substitute in a spider one more time and simplify.



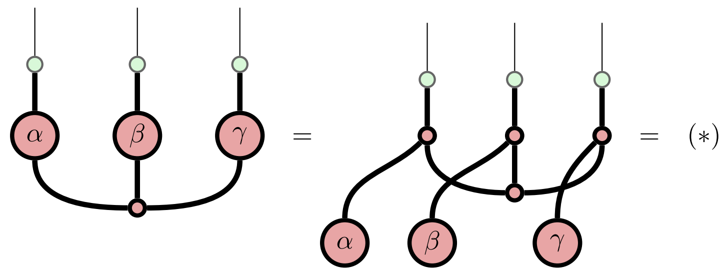
$\square$

Lemma 3.6. [7] *For arbitrary phases* α, β, γ

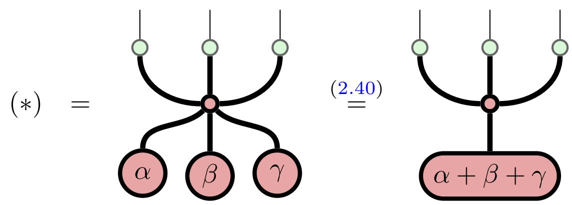


[7]For the rest of this paper, green and red denote the Z and X bases, respectively.

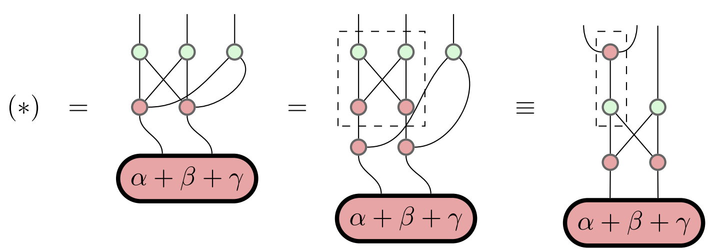
Proof. Recalling Definition 2.38, we have the following.



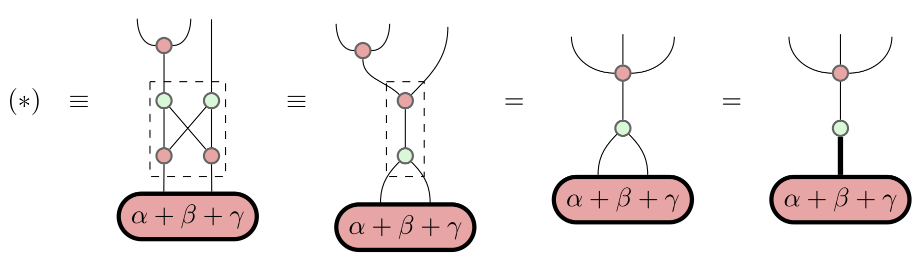
We simplify to one X spider via Theorem 2.17 and then add the phases together.



We rewrite the quantum (doubled) spider as double classical spiders, which we expand via Theorem 2.17. Then, we apply the rule from Definition 2.46 (indicated by the dashed lines).



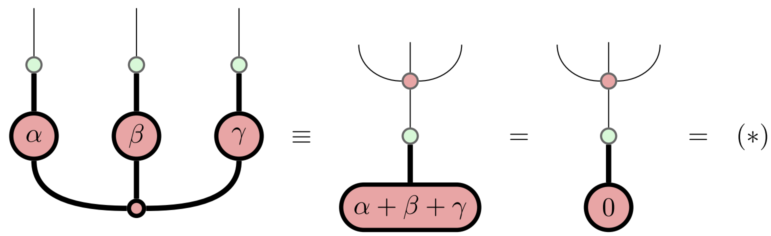
We apply the rule from Definition 2.46 again. Finally, we simplify to one X spider via Theorem 2.17 and rewrite the quantum wire.



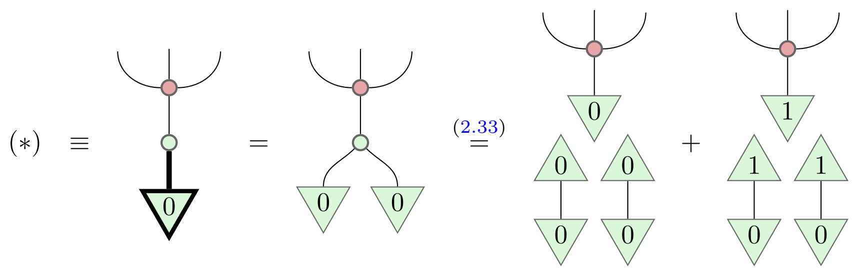
□

Lemma 3.7. *We consider the scenario from Lemma 3.6. If $\alpha + \beta + \gamma = 0$, then the parity ($\mathbb{Z}_2$ sum) of the diagram is even (0).*

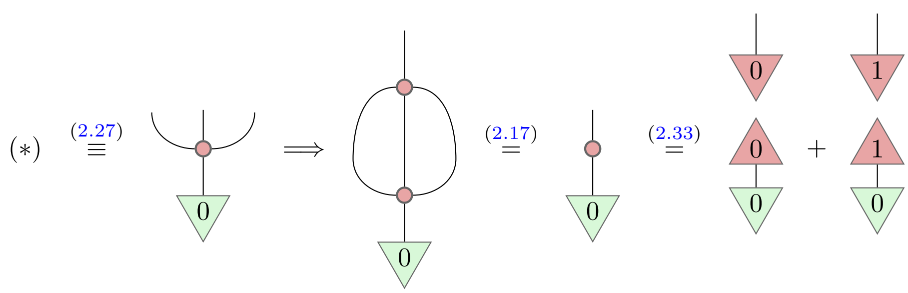
Proof. From Lemma 3.6, we have



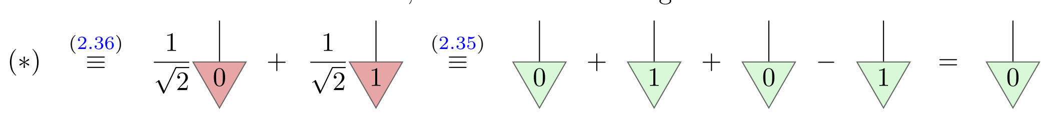
Recalling Remark 2.41, we have



This simplifies according to the definition of ONBs. We compute the parity by using an X spider to combine the three outputs.



From the definition of the X basis, we have the following.



Thus, the parity is even (0). □

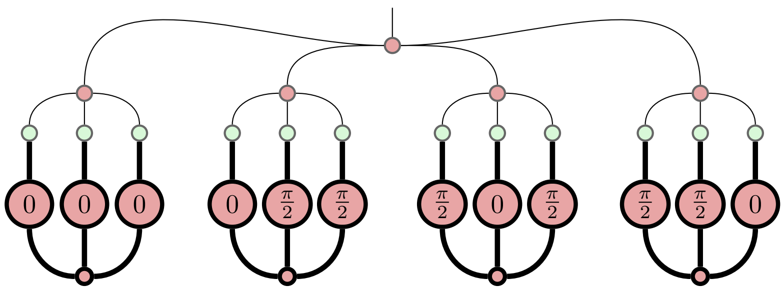
Lemma 3.8. *We consider the scenario from Lemma 3.6. If $\alpha + \beta + \gamma = \pi$, then the parity ($\mathbb{Z}_2$ sum) of the diagram is odd (1).*

Proof. The proof is given by the exact same steps as above. □

With all preliminaries taken care of, we proceed to prove Theorem 3.1.

Proof. Suppose that Alice, Bob, and Charlie are sufficiently separated such that there is no causal arrow between any of them. Consider the following experiment. Dave—who has a causal arrow to and from each party—prepares three qubits in a GHZ state. He then sends one qubit each to Alice, Bob, and Charlie, along with directions to apply a phase of either 0 or $\frac{\pi}{2}$ in the X basis. Alice, Bob, and Charlie will then measure their qubits in the Z basis and communicate their results to Dave. A phase of 0 in the X basis followed by a measurement in the Z basis is termed a Z *measurement*, and a phase of $\frac{\pi}{2}$ in the X basis followed by a measurement in the Z basis is termed a Y *measurement*. Dave will ensure that one of the following measurement combinations occurs: ZZZ, ZYY, YZY, or YYZ.

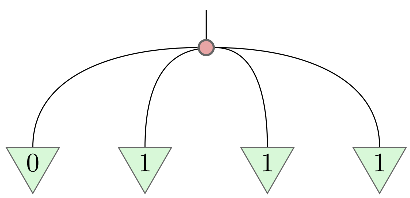
The essence of this proof is to show that it is impossible for a local hidden variables theory to store measurement outcomes for these four scenarios in advance in a way that is consistent with quantum mechanics. We do this by computing the overall parity of these four possible scenarios.



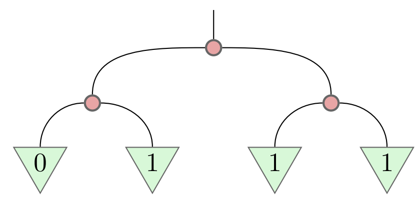
$= \quad (*)$

By Lemmas 3.7 and 3.8,

$(*) \quad \equiv$

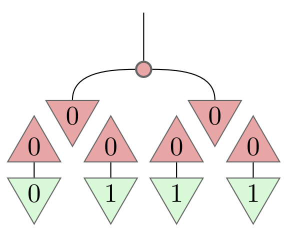


$\overset{(2.17)}{=}$

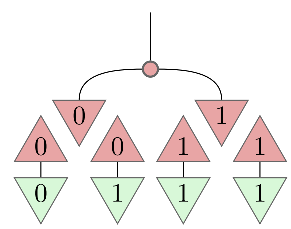


Then,

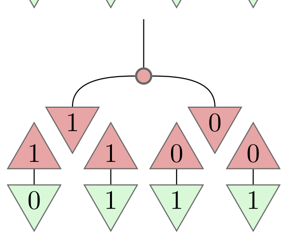
$(*) \quad \overset{(2.33)}{\equiv}$



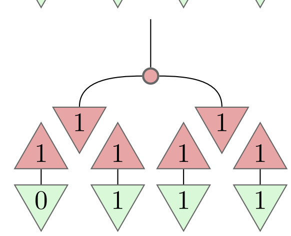
$+$



$+$

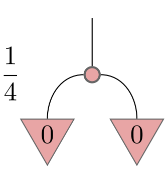


$+$

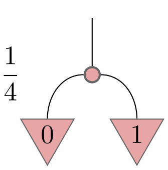


Recalling the relationship between the Z and X basis, we have the following

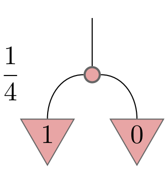
$(*) \quad \overset{(2.36)}{\equiv} \quad \frac{1}{4}$



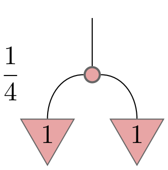
$+ \quad \frac{1}{4}$



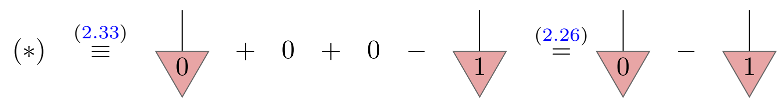
$- \quad \frac{1}{4}$



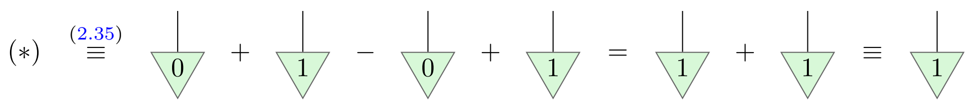
$- \quad \frac{1}{4}$



From (2.33), we have

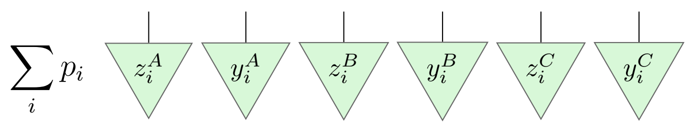


Finally, from the definition of the X basis, we have

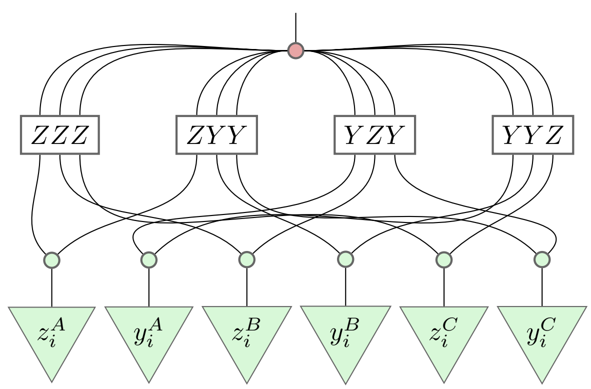


Thus, the overall parity is odd. We now show that computing the overall parity with local hidden variables yields the opposite result.

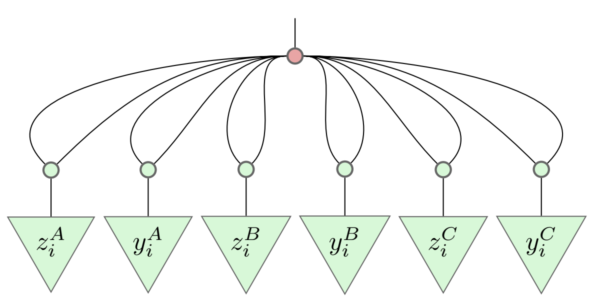
In a local hidden variables theory, the values for any possible measurement are determined by variables that exist before the measurement. Each qubit faces two possible measurements: a Z measurement or Y measurement. Thus, we need two variables for each qubit. Moreover, these variables could be probabilistic, so we have the following probability distribution.



We proceed to show that *any* state in this probability distribution yields a contradictory result. We fix i and compute the overall parity of the four measurement possibilities for this particular state. Note that the four boxes in this diagram are merely labels denoting the four possible scenarios, which explain the number of wires coming from each variable.

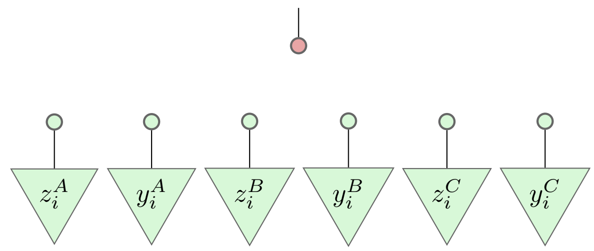


Each Z spider sends exactly two outputs to the X spider, so we simplify our drawing of the diagram.

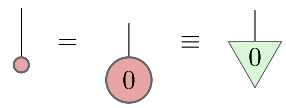


$= \quad (*)$

By Definition 2.45,

$(*) \equiv$ 

By causality, Property 2.52, we can discard all of the states. And recalling Remark 2.41, we have the following.

$(*) \equiv$ 

Thus, the overall parity is even, which contradicts the result from categorical quantum mechanics. We conclude that any local hidden variables theory is inconsistent with quantum mechanics. □

4. Discussion

It may not be immediately obvious why *locality* was the cause of the contradiction; why does this proof not refute hidden variables theories in general? If we allow our hidden variables theory to be non-local, then we can have causal arrows going both ways between Alice, Bob, and Charlie, so each measurement outcome is informed by the other two. Then, we can use four global variables—one for each measurement scenario—and ensure that the parity of each scenario agrees with the calculation done with categorical quantum mechanics. This demonstrates how non-locality is a special case of what is called *contextuality*, which is the idea that hidden variables theories only work if they can store multiple measurement outcomes in one variable.

After establishing non-locality in categorical quantum mechanics, it is then natural to ask what causes such behavior. Coecke et al. argue in [3] that non-locality in GHZ-Mermin scenarios (such as the one in our proof) necessitate strongly complementary ONBs. This is not surprising to us, as strong complementarity was necessary in our proof to compute the parity with the local hidden variables theory. It also turns out that the structure of the phase group is essential. For example, one can restrict the phase group to four elements $\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ and obtain a group isomorphic to $\mathbb{Z}_4$—and this still exhibits non-locality. In [2], Coecke and Kissinger demonstrate how one can construct a seemingly quantum theory with a phase group of $\mathbb{Z}_2 \times \mathbb{Z}_2$ that exhibits quantum behavior like strong complementarity and non-separability but does not exhibit non-locality.

Acknowledgments

Special thanks to my mentor, Yanbo Chen, for his support and advice throughout this process. Thanks as well to Peter May and everyone else who makes this incredible program a reality.

5. Bibliography

Many of the categorical definitions are based on those from [1]. Other categorical notions and the quantum theory is based on [2]. The proof of Theorem 3.2 is a generalization of a proof presented in [2]. Depicting non-locality in categorical quantum mechanics via GHZ-Mermin scenarios is done in [3] and using the same scenario we present in [2]. All such arguments, including the one presented in this paper, are based on [4]. Of course, the first proof of Bell's theorem—given in a very different manner—is [5].

References

[1] Brendan Fong and David I. Spivak. An Invitation to Applied Category Theory. Cambridge University Press. 2019.

[2] Bob Coecke and Aleks Kissinger. Picturing Quantum Processes. Cambridge University Press. 2017.

[3] Bob Coecke, Ross Duncan, Aleks Kissinger, Quanlong Wang. Strong Complementarity and Non-locality in Categorical Quantum Mechanics. arXiv:1203.4988

[4] N. David Mermin. Quantum mysteries revisited. American Journal of Physics. 1990.

[5] J.S. Bell. On the Einstein Podolsky Rosen Paradox. https://cds.cern.ch/record/111654/files/vol1p195-200_001.pdf