Generalised Compositional Theories and Diagrammatic Reasoning

Bob Coecke, Ross Duncan, Aleks Kissinger, and Quanlong Wang

Bob Coecke
University of Oxford, Department of Computer Science,
Wolfson Building, Parks Road, Oxford OX1 3QD, UK
e-mail: coecke@cs.ox.ac.uk

Ross Duncan
University of Strathclyde, Department of Computer and Information Sciences,
Livingston Tower, 26 Richmond Street, Glasgow G1 1XH, UK
e-mail: ross.duncan@strath.ac.uk

Aleks Kissinger
University of Oxford, Department of Computer Science,
Wolfson Building, Parks Road, Oxford OX1 3QD, UK
e-mail: alek@cs.ox.ac.uk

Quanlong Wang
Beihang University, School of Mathematics and System Sciences,
XueYuan Road No.37, HaiDian District, Beijing, China
e-mail: qlwang@buaa.edu.cn

1

This chapter provides an introduction to the use of *diagrammatic language*, or perhaps more accurately, *diagrammatic calculus*, in quantum information and quantum foundations. We illustrate the use of diagrammatic calculus in one particular case, namely the study of *complementarity* and *non-locality*, two fundamental concepts of quantum theory whose relationship we explore in later part of this chapter.

The diagrammatic calculus that we are concerned with here is not merely an illustrative tool, but it has both (i) a conceptual physical backbone, which allows it to act as a foundation for diverse physical theories, and (ii) a genuine mathematical underpinning, permitting one to relate it to standard mathematical structures.

(i) The conceptual physical backbone concerns *compositionality*. Given two systems, there is also a composite system. This notion of composition is a primitive ingredient of the diagrammatic language. Moreover, the basic elements of the diagrammatic language are *processes*, and states are identified with preparation processes. This paves the way for a framework of *generalised compositional theories* (GCTs), named in analogy to generalised probabilistic theories [1]. The latter have recently received much attention because one can better understand a theory—quantum theory in particular—by studying it as merely a member of a broader class of theories. Notably, the study of non-locality within this framework has provided important new insights [2, 3]. Whereas generalised probabilistic theories discard everything except the convex probabilistic structure, in contrast, GCTs focus on composition. This approach is informed by techniques used in computer science, logic, and the branch of mathematics called category theory, however its roots can be traced to Schrödinger's conviction that the essential characteristic of quantum theory is the manner in which systems compose [4].

(ii) On the other hand, the diagrammatic language has a well-defined mathematical meaning, which permits any diagram to be interpreted as a definite object in various other concrete mathematical models, for example in Hilbert spaces. This translation can be carried out in a formally precise manner, so that reasoning in the diagrammatic calculus produces true equations in the chosen model. At the same time, the relationship between what is provable in the calculus and what is provable in concrete models can be described to a high degree of precision.

We won't discuss this mathematical basis in detail here, however it may be summarised as follows: the diagrammatic calculus is itself a GCT, and GCTs form a certain class of *monoidal categories*, also known as *tensor categories*. The use of diagrammatic languages for tensors traces back to Penrose in the early 1970's [5], but was only placed on a formal mathematical basis in the late 1980's [6, 7]. Their use in quantum foundations and quantum information began with an abstract (partial) axiomatisation of Hilbert spaces in terms of these categories [8], eventually resulting in so-called *quantum picturalism* [9]. Meanwhile, the diagrammatic compositional language has been adopted by several researchers in quantum foundations [10, 11]. The particular developments related here been used to solve problems in quantum foundations [12, 13] and quantum computation [14, 15, 16].

1 Introduction to Quantum Picturalism

1.1 Theories and Diagrams

A generalised compositional theory consists of *systems*, or more accurately *types of systems*, and *processes* which transform systems. A process f which transforms systems of type A into systems of type B is written $f : A \to B$. At the highest level of generality we do not need to give any details as to what A, B, or f are: it is enough to know that that f accepts systems of type A as inputs and produces systems of type B as outputs. The important thing is how systems and processes are combined.
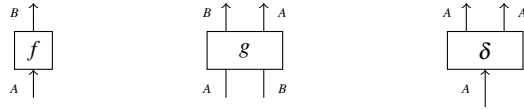
Mathematically speaking, general compositional theories are *strict symmetric monoidal categories*, and a full exposition of their properties would require a lengthy detour into category theory. The interested reader can refer to Mac Lane's classic text [17] for a thorough treatment. However, we can avoid reading Mac Lane's book[1] by adopting a diagrammatic notation, which absorbs all of the relevant equations into the syntax. This notation is the subject of the first section of this paper.

We will represent processes by *diagrams*, consisting of boxes and wires. The wires are labelled by systems, and the boxes by basic processes[2]. Wires join boxes at the top and bottom; the wires below correspond to the input systems of the process,

---

[1] We jest; reading Mac Lane's book is eventually unavoidable, however the paper [18] is an easy introduction to the subject of monoidal categories.

[2] The term "basic" simply means a process whose internal structure is of no interest. Typically we construct diagrams from some given set of basic processes.

and those at the top correspond to the output systems. For example:
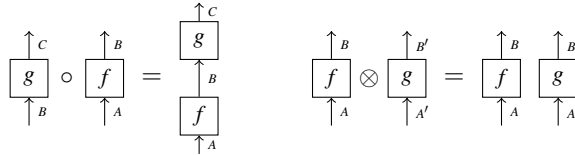
$$f : A \to B \qquad g : A \otimes B \to B \otimes A \qquad \delta : A \to A \otimes A$$

The same is true for the diagram as a whole: the wires entering the bottom of the diagram are its input systems, and those leaving from the top are its outputs.

Given processes $f : A \to B$ and $g : B \to C$, it seems obvious that doing f then g is again a process, and we write $g \circ f : A \to C$ to denote this process. In other words, processes admit *sequential combination*; we will usually call this operation *composition*.

Similarly, a pair of systems, say A and B, can be taken together and viewed as a single system, $A \otimes B$. Now, given a pair of processes $f : A \to B$ and $g : A' \to B'$, a new process is obtained by placing them in parallel. We denote the combined process $f \otimes g : A \otimes A' \to B \otimes B'$. This operation of parallel combination is called *tensor*.

In the diagrammatic notation, composition is expressed by plugging the outputs of one box into the inputs of another, and the tensor is given by juxtaposition.

We require that both operations, composition and tensor, are associative and obey the interchange law,

$$(f \otimes g) \circ (h \otimes k) = (f \circ h) \otimes (g \circ k) . \qquad (1)$$

In the graphical notation, all of these equations become trivial: they boil down the statement that the three diagrams below are unambiguous.
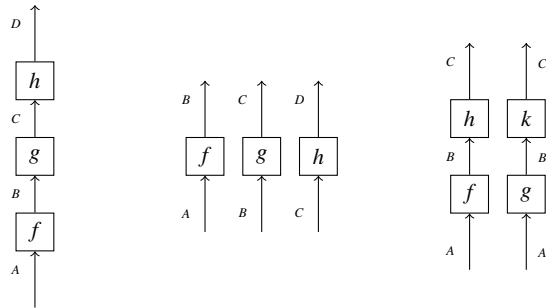
While it is easy to translate these diagrams back into conventional notation, to do so we must make a *choice* of where to put the brackets, even though the theory

tells us this choice does not matter. This highlights a key advantage of working with diagrams, namely that the objects which are *equal* in the theory produce the *same diagram*.

In addition to the two operations, composition and tensor, every generalised compositional theory is equipped with certain primitive processes. The simplest process is the process which doesn't do anything at all, simply returning unchanged the system given to it. We assume that for every system A such a null process, called the *identity* and written $1_A : A \rightarrow A$, exists. The fact that it does nothing is expressed by the equations

$$1_B \circ f = f = f \circ 1_A$$

for all processes $f : A \rightarrow B$. The identity process $1_A : A \rightarrow A$ is drawn as a wire without any box on it, while the identity for $A \otimes A'$ is simply the tensor product $1_A \otimes 1_{A'}$, i.e. two wires.

$$1_A = \; {}_A \Big\uparrow \qquad 1_{A \otimes A'} = \; {}_A \Big\uparrow \; \otimes \; {}_{A'} \Big\uparrow \quad = {}_A \Big\uparrow \; {}_{A'} \Big\uparrow$$

Once again we see an equation absorbed into the notation: since the identity has no effect on a process, the *length* of the wires attached to a box makes no difference.
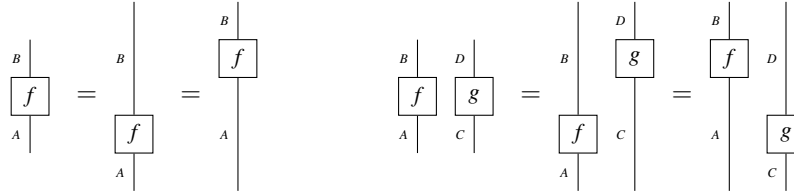
In addition, for every pair of systems A and B there is a process $\sigma_{A,B} : A \otimes B \rightarrow B \otimes A$ which exchanges the two systems. The class of theories we consider here are *symmetric*: swapping two systems twice has no effect, hence the equation

$$\sigma_{B,A} \circ \sigma_{A,B} = 1_{A \otimes B}$$

holds for all systems A and B Graphically, the swap is just the crossing of two wires:
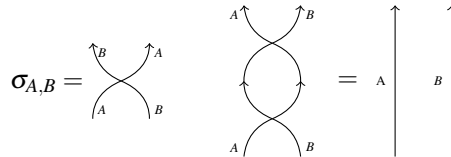
In fact, the swap should satisfy some further *coherence equations*, the details of which can be found in [17]. However, we can again make the graphical notation do the work by allowing wires to cross freely in the diagrams, and saying that only the connectivity of the wires matters, and not their configuration in the page. For example, the following diagrams are equal:
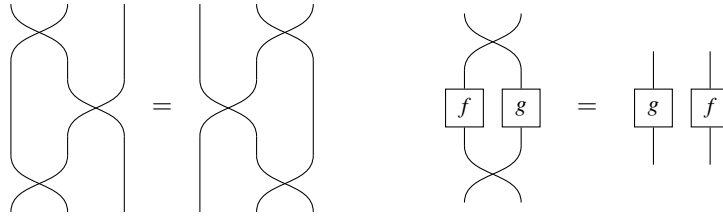
Note that we do not distinguish between wires crossing over and crossing under.

A process may produce an output without having to consume an input first, or vice versa. Therefore we introduce a null system, or empty system, which we denote I. Hence a process that produces an A from nothing would be written $p : I \to A$. Like the identity process, the null system obeys some equations:

$$A \otimes I = A = I \otimes A \qquad \text{and} \qquad 1_I \otimes f = f = f \otimes 1_I,$$

for all systems A and all processes f. As suggested by the preceding equations, I is represented as empty space in the diagram, and its identity process $1_I : I \to I$ is represented by the empty diagram.

A process of type $s : I \to I$ is called a *scalar*; this name will be justified later. It is clear from the diagrammatic notation that given scalars s and s' we have $s \circ s = s \otimes s' = s' \otimes s = s' \circ s$; i.e. the scalars form a commutative monoid[3].

In the preceding text we have introduced various transformations of diagrams which, we claim, do not change anything. It is reasonable to ask: when are two diagrams considered to be equal? We use a very intuitive notion here: **Two diagrams are considered equal when, keeping the inputs and outputs fixed, one may be transformed to the other by purely topological transformations.** In other words, if starting from one diagram we—by crossing or uncrossing wires, stretching wires, moving boxes along wires, translating boxes in the plane (while maintaining their connections), etc—arrive at the other, then they are equal. In particular, since scalars
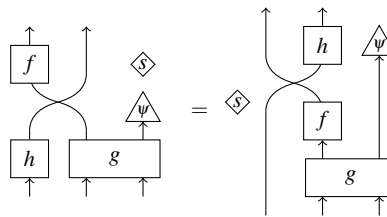
Fig. 1 Examples of topologically equivalent diagrams.

---

[3] This is true even for non-symmetric monoidal categories; see [7].

are not connected to the inputs or outputs of the diagram, they may be placed any-where in the diagram without altering its meaning.

Example 1. The simplest non-trivial example is the theory with one primitive sys-tem, denoted *u*, and whose processes are generated by the identity and swap. We call this theory **SymGrp**. Since there is only one basic system, every other system is just an *n*-fold tensor power of *u*, hence the systems of the theory can be identified with the natural numbers. In this theory, a process $p : n \to n$ is nothing more than a sequence of swaps; i.e. a permutation on the *n*-element set. Hence **SymGrp** is exactly the theory of the symmetric groups.
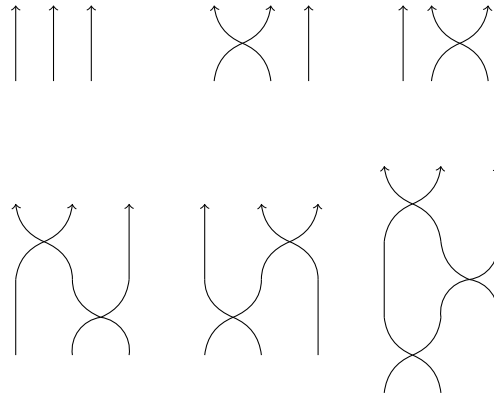

Fig. 2 Example: the symmetric group S_3 presented as diagrams.

Example 2 (Finite-dimensional Hilbert spaces). The theory called **FHilb** has as its systems all finite-dimensional complex Hilbert spaces. The processes of this theory are all linear maps $f : A \to B$. The sequential composition of processes is the usual composition of linear maps, and the tensor is the usual Kronecker product of vector spaces and maps. The identity process is the identity map, the swap is the evident permutation map, and the null system is the base field, $\mathbb{C}$. Since a linear map $\mathbb{C} \to \mathbb{C}$ is totally determined by its value at 1, we see that the scalars of **FHilb** are nothing more than the complex numbers themselves.

We write **FHilb**$_D$ to denote the subtheory **FHilb** restricted to Hilbert spaces of dimension D^n and linear maps between them, for some fixed D. For convenience, we refer to **FHilb**$_2$ as **Qubit**. Notice that the systems of **Qubit** are all tensor powers of $\mathbb{C}^2$, and its processes include all quantum circuits, state preparations, and post-selected measurements, justifying the name.

Remark 1. Note that we must specify what the tensor product is to specify what the theory is. For example, another equally valid theory is the collection of finite dimensional Hilbert spaces and linear maps, but with the direct sum as the tensor.

This is again a general compositional theory, although since it lacks certain other features we will require later, it will play no role in this presentation.

Example 3 (Sets and Relations). An example with very different flavour, but most of the same structure is **FRel**. The systems of **FRel** are all finite sets (considered up to isomorphism[4]), and the processes $r : X \to Y$ are relations between X and Y, that is subsets of $X \times Y$. The composition of relations is given by

$$s \circ r = \{(x,z) \mid \exists y \text{ s.t. } (x,y) \in r \text{ and } (y,z) \in s\}.$$

The identity process is the diagonal relation,

$$1_X = \{(x,x) \mid x \in X\}.$$

The tensor product in **FRel** is the cartesian product $X \otimes Y = X \times Y$, which takes the form

$$r \otimes r' = \{((x,x'),(y,y')) \mid (x,y) \in r \text{ and } (x',y') \in r'\}$$

on processes. The null system is the singleton set $\{*\}$, for which we have $\{*\} \times X \cong X$ for all sets X. There are exactly two relations from $\{*\}$ to itself, namely the total relation and the empty relation. Hence, the scalars of **FRel** are the Boolean monoid, i.e. $\mathbb{Z}_2$ with the usual multiplication.

An important subtheory of **FRel** is **FSet**, obtained by restricting the to relations which are functions: that is, relations $r : X \to Y$ where each x is related to exactly one y. Just as in the case of **FHilb**, we can consider restrictions of **FRel** to systems generated by a set of size D, which we call **FRel**$_D$. For example, **FRel**$_2$ contains all the Boolean functions. The intersection of **FRel**$_2$ and **FSet** consists of precisely the Boolean functions; this theory we denote **Bool**. Many other interesting theories are subtheories of **FRel**; we'll meet some more later.

Since generalised compositional theories all share certain basic structure, it is natural to consider maps between them. Given two such theories **C** and **D**, a map $F : \mathbf{C} \to \mathbf{D}$ consists of an assignment of each system A in **C** to a system FA in **D**, and an assignment of each process $f : A \to B$ in **C** to a process $Ff : FA \to FB$ in **D**, obeying the following equations:

$$F(A \otimes B) = FA \otimes FB \qquad\qquad FI = I$$
$$F(g \circ f) = Fg \circ Ff \qquad\qquad F(f \otimes g) = Ff \otimes Fg$$
$$F1_A = 1_{FA} \qquad\qquad F\sigma_{A,B} = \sigma_{FA,FB}$$

In the mathematics literature, such a map is called a *strict symmetric monoidal functor*; again, see Mac Lane [17] for the details. The important point to note is that the mapping F sends wires to wires. Therefore, to specify such a mapping it is enough to specify the image of the boxes in a diagram, ensuring that composition and tensor are respected.

---

[4] Since we identify sets of the same cardinality, we can equivalently say that the systems of **FRel** are just the natural numbers.

Example 4. We can define a map $R_D : \mathbf{SymGrp} \to \mathbf{FHilb}$ by setting $R_D(u) = \mathbb{C}^D$ and then everything else is defined by the requirement that R_D is a strict symmetric monoidal functor. Thus we have a D^n dimensional representation of the symmetric group S_n for every D.

In fact, this construction applies equally well to any generalised compositional theory $\mathbf{C}$: all that is required is an assignment of the unique primitive system u to some system of $\mathbf{C}$. Therefore every generalised compositional theory contains all the symmetric groups.

Given a mapping between theories it is easy to calculate the image of a given diagram. One must recursively partition the diagram into tensors and compositions of smaller diagrams until each partition contains exactly one element—that is, either a single wire, a crossing of wires, or a box. The interchange law (Equation 1) guarantees that the result does not depend on the partition chosen.
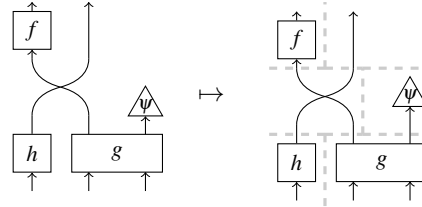

We may now state:

Theorem 1 (Fundamental Theorem of Diagrams). *Given any two generalised compositional theories* $\mathbf{C}$ *and* $\mathbf{D}$*, and a map* $F : \mathbf{C} \to \mathbf{D}$*, for any two diagrams d and d' in* $\mathbf{C}$*, if d = d' as diagrams then Fd = Fd' in* $\mathbf{D}$*.*

This theorem has many variations, and we refer the reader to Selinger's survey article [19] for the full details.

Remark 2. In the diagrams to come, we will often use horizontal separation to indicate separation in space and vertical separation to indicate separation in time. For example,
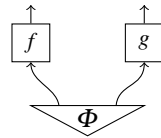

depicts the creation of two systems by the process Φ, which then become spatially separated over some time and are acted upon by processes f and g respectively. Since, as we already know, topologically equivalent diagrams are equal, these separations have no formal status and are purely illustrative.

1.2 Rewrites and Models

Since we wish to generalise over many concrete mathematical structures, we are particularly interested in theories which can be specified *axiomatically*. That is, to specify the theory we state (i) the list of basic systems—typically we'll only have one basic system, the rest being generated by the tensor product—and (ii) the basic processes. The processes of the theory are then *all* the diagrams which can be constructed from these processes and nothing else.

Example 5 (Boolean Circuits). A simple example of a compositional theory is **BoolCirc**, the theory of boolean circuits. This theory has only one basic system, the bit b, and the basic processes are the logic gates:

$$\wedge : b \otimes b \to b \qquad \vee : b \otimes b \to b \qquad \neg : b \to b \qquad \textbf{FAN} : b \to b \otimes b$$
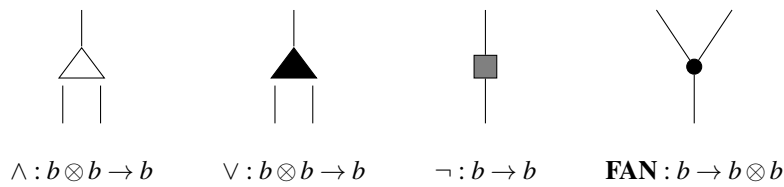
A process in this theory is a circuit for computing some boolean function, built up from these basic gates.
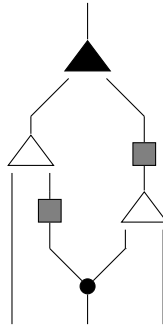

Fig. 3 A Boolean circuit to compute $(x \wedge \neg y) \vee \neg(y \wedge z)$.

It is tempting to assume that **BoolCirc** is related to the theory of Boolean functions, and we can make this precise by specifying a mapping $B : \textbf{BoolCirc} \to \textbf{Bool}$. We assign $B(b) = \{0, 1\}$ and define B on the basic processes as follows:

$$B(\wedge) = a : \begin{cases} 00 \mapsto 0 \\ 01 \mapsto 0 \\ 10 \mapsto 0 \\ 11 \mapsto 1 \end{cases} \qquad B(\vee) = o : \begin{cases} 00 \mapsto 0 \\ 01 \mapsto 1 \\ 10 \mapsto 1 \\ 11 \mapsto 1 \end{cases}$$

$$B(\neg) = n : \begin{cases} 0 \mapsto 1 \\ 1 \mapsto 0 \end{cases} \qquad B(\textbf{FAN}) = \delta : \begin{cases} 0 \mapsto 00 \\ 1 \mapsto 11 \end{cases}$$

The mapping B assigns to each diagram the boolean function normally associated with it. However this is not the only possibility. Consider the following mapping, $P : \textbf{BoolCirc} \to \textbf{Bool}$. Once again $P(b) = \{0,1\}$, but now we have the following assignment of processes:

$$P(\wedge) = a : \begin{cases} 00 \mapsto 0 \\ 01 \mapsto 0 \\ 10 \mapsto 0 \\ 11 \mapsto 1 \end{cases} \qquad P(\vee) = p : \begin{cases} 00 \mapsto 0 \\ 01 \mapsto 1 \\ 10 \mapsto 1 \\ 11 \mapsto 0 \end{cases}$$

$$P(\neg) = i : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 1 \end{cases} \qquad P(\textbf{FAN}) = \delta : \begin{cases} 0 \mapsto 00 \\ 1 \mapsto 11 \end{cases}$$

The mapping P assigns to each $d : b^n \to b$ in **BoolCirc** an n-variable polynomial over the ring $\mathbb{Z}_2$. (More generally a circuit with multiple outputs produces a list of polynomials, one for each output.)

In fact, as the example of P suggests, the diagrams of **BoolCirc** admit an interpretation in any setting with two binary operations and one unary operation. This is not entirely satisfactory. In order to capture more than the bare syntax of any given theory we need to impose some additional equations on the class of diagrams. We do this via *rewrite rules*.

A rewrite rule consists of a pair of diagrams of the same type, for example $d : A \to B$ and $d' : A \to B$. If this rule is called r then we write $r : d \Rightarrow d'$, or diagrammatically
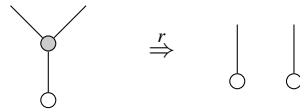
Whenever d occurs as a subdiagram of a larger diagram e then we can replace d with d' in e, written $e[d] \overset{r}{\Rightarrow} e[d']$, or in diagrams:
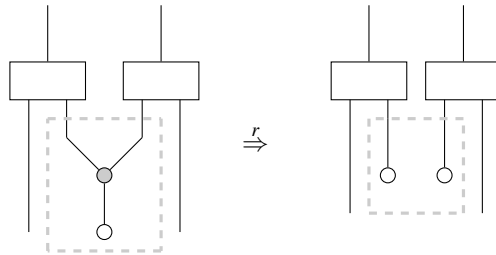
Rewrite rules allow us to define a notion of equality in addition to the basic equality of diagrams. Given a collection of rewrite rules $\mathscr{R}$ we write $d \overset{\mathscr{R}}{\Rightarrow} d'$ if there is some rewrite sequence in $\mathscr{R}$ taking d to d'. Evidently $\overset{\mathscr{R}}{\Rightarrow}$ is a transitive relation; let $\overset{\mathscr{R}}{=}$ be its symmetric, reflexive closure. Then we say that two processes are equal according to $\mathscr{R}$ if their corresponding diagrams satisfy $d \overset{\mathscr{R}}{=} d'$. Typically we'll exhibit this equivalence as a sequence of rewrites.

Example 6 (Boolean circuits). Consider the following two rewrite rules for **BoolCirc**, expressing respectively the distributivity of AND over OR, and (one half of) De Morgan's law.
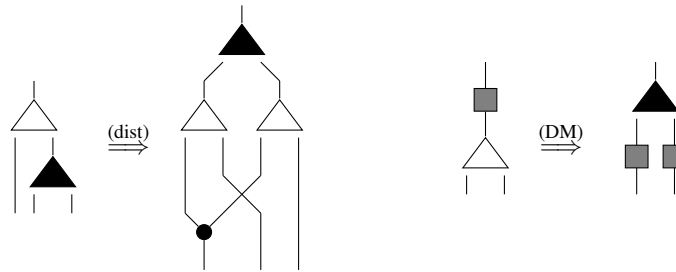
Now we can show that a certain Boolean circuit can be transformed into its disjunctive normal form:
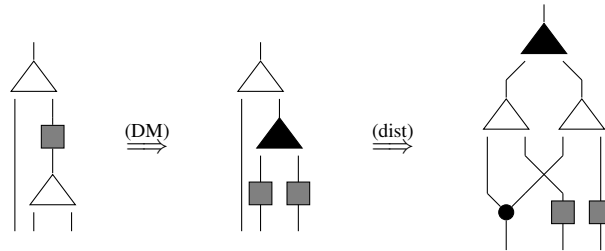
Given a theory **C**, a set of rewrite rules $\mathscr{R}$, and a mapping $F : \mathbf{C} \to \mathbf{D}$, we can ask the following question: if $d \overset{\mathscr{R}}{=} d'$ in **C**, is it the case that $Fd = Fd'$ in **D**?
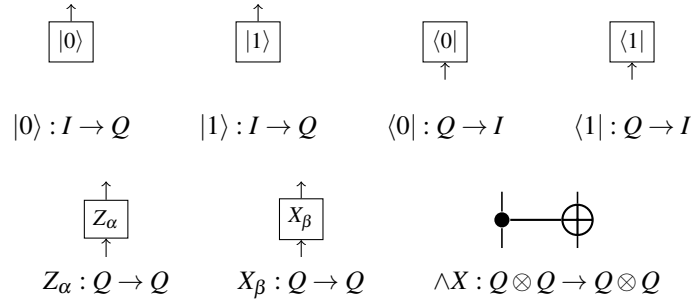
This property is called *soundness*. A sound mapping $F : \mathbf{C} \to \mathbf{D}$ is called an *interpretation* of **C** in **D**, and the image of **C** in **D** is called a *model*. In the example above, the mapping B is sound, hence it provides an interpretation of **BoolCirc** (and $\mathscr{R}$) in **Bool**; on the other hand P does not, due to the failure of De Morgan's law. Generally speaking we will always work with a given set of rewrite rules and a given interpretation map, so we will usually say "*the* **D** interpretation of **C**", although in principle there could be many.

Remark 3. The converse property to soundness, $Fd = Fd'$ implies $d = d'$, is called *completeness*. An interpretation which is both sound and complete provides an iso-morphism between the formally presented theory and its model. While checking

soundness is straightforward, showing completeness is often much more difficult[5]. On the other hand, not having completeness means there are multiple models of a given theory, and the study of the differences between such models is often informative.

Before moving on, we'll introduce an important example, and its standard model.

Example 7 (Quantum Circuits). Similar to the example of Boolean circuits, we can also view (post-selected) quantum circuits as generalised compositional theory, called **QuCirc**. Again we have a single basic system, the qubit Q, and the basic processes are a collection of unitary gates, state preparations, and projections from which we construct the other quantum circuits.

$$|0\rangle : I \to Q \qquad |1\rangle : I \to Q \qquad \langle 0| : Q \to I \qquad \langle 1| : Q \to I$$

$$Z_\alpha : Q \to Q \qquad X_\beta : Q \to Q \qquad \wedge X : Q \otimes Q \to Q \otimes Q$$

From these basic elements we can write down any quantum circuit. We now define the standard interpretation of **QuCirc** into **Qubit**.

$$\llbracket Q \rrbracket = \mathbb{C}^2$$

$$\left\llbracket \boxed{|0\rangle} \right\rrbracket = |0\rangle \qquad \left\llbracket \boxed{|1\rangle} \right\rrbracket = |1\rangle \qquad \left\llbracket \boxed{\langle 0|} \right\rrbracket = \langle 0| \qquad \left\llbracket \boxed{\langle 1|} \right\rrbracket = \langle 1|$$

$$\left\llbracket \boxed{Z_\alpha} \right\rrbracket = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix} \qquad \left\llbracket \boxed{X_\beta} \right\rrbracket = \begin{pmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}$$

$$\left\llbracket \; \bullet\!\!-\!\!\oplus \; \right\rrbracket = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

---

[5] To show completeness for a rewrite theory it is typically necessary, but rarely sufficient, to check that the rewrite rules are *confluent*; that is, whenever two rewrites simultaneously apply to a given diagram, then the choice between then (eventually) does not matter. Since this property must hold for every diagram and every pair of rewrites, even a simple rewrite system can produce an extremely large number of cases, necessitating a computer-assisted proof. For example see the work of Lafont on Boolean circuits [20].

Thanks to the well-known universality result [21] this interpretation demonstrates that **QuCirc** can represent all unitary maps between qubits. In fact, since we have the projections $\langle 0|, \langle 1|$, all linear maps can be represented. Note, however, that although all quantum circuits can be represented, without a set of rewrite rules **QuCirc** cannot express any non-trivial equalities between them. We could propose various sound equations here, but there is no known collection of rewrite rules which makes **QuCirc** complete with respect to this interpretation into **Qubit**. If such a set of rewrites did exist, it would constitute provide a presentation of the unitary group by generators and relations.

1.3 The Dagger

Now we introduce the *dagger*. This is simply an operation on the processes of a theory, sending every process $f : A \to B$ to another process $f^\dagger : B \to A$. We call $f^\dagger$ the *adjoint* of f. In the graphical calculus, we represent the dagger by a flip in the horizontal axis:

$$\left(\boxed{f \, /} \right)^\dagger = \boxed{f \, \backslash}$$

Note that we have made the box asymmetric to make this flipping evident. For more general diagrams, the dagger flips a diagram upside down, preserving all the internal structure. Taking this claim at face value, we can derive the key properties of the dagger:

$$(f^\dagger)^\dagger = f$$

$$(g \circ f)^\dagger = f^\dagger \circ g^\dagger$$

$$(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$$

$$1_A^\dagger = 1_A$$

$$\sigma_{A,B}^\dagger = \sigma_{B,A}$$

The dagger allows two important concepts to be defined.

Definition 1. A process $f : A \to B$ is called *unitary* if $f \circ f^\dagger = 1_B$ and $f^\dagger \circ f = 1_A$. A process is called *self-adjoint* when $f^\dagger = f$.

Example 8 (Finite-dimensional Hilbert spaces). The theory **FHilb** admits a dagger: it is the usual adjoint of a linear map. In this theory, the abstract definitions of unitarity and self-adjointness coincide with the usual one.

Example 9. In the theory **FRel**, the dagger of a relation $r : X \to Y$ is defined by the converse relation, i.e.

$$r^\dagger = \{(y,x) \mid (x,y) \in r\}$$

Here, unitary processes are exactly those relations which encode permutations. A relation is self-adjoint whenever it is symmetric. Hence the self-adjoint unitaries in **FRel** are exactly the permutations of order 2.

We extend the definition of mapping to demand that it also preserves the dagger. That is, given two theories with dagger, we require that a map $F : \mathbf{C} \to \mathbf{D}$ satisfies

$$F(f^\dagger) = (Ff)^\dagger$$

Example 10 (Quantum Circuits). We define a dagger on **QuCirc** as follows:

$$\left(\boxed{|0\rangle} \right)^{\dagger} = \boxed{\langle 0|} \qquad\qquad \left(\boxed{|1\rangle} \right)^{\dagger} = \boxed{\langle 1|}$$

$$\left(\boxed{\langle 0|} \right)^{\dagger} = \boxed{|0\rangle} \qquad\qquad \left(\boxed{\langle 1|} \right)^{\dagger} = \boxed{|1\rangle}$$

$$\left(\boxed{X_\alpha} \right)^{\dagger} = \boxed{X_{-\alpha}} \qquad\qquad \left(\boxed{Z_\beta} \right)^{\dagger} = \boxed{Z_{-\beta}}$$

$$\left(\bullet\!\!-\!\!-\!\!\oplus \right)^{\dagger} = \bullet\!\!-\!\!-\!\!\oplus$$

It's now easy to check that the interpretation map introduced earlier, $\llbracket \cdot \rrbracket : \mathbf{QuCirc} \to \mathbf{Qubit}$ preserves the dagger as required.

Remark 4. The theory of Boolean circuits, **BoolCirc**, does not admit a dagger. However, we could formally add new basic processes corresponding to the adjoints of the basic processes of **BoolCirc** and thus define a new theory, $\mathbf{BoolCirc}^{\dagger}$. Since the converse of a function is not in general a function, the interpretation $B : \mathbf{BoolCirc} \to \mathbf{Bool}$ no longer makes sense. Instead we must interpret $\mathbf{BoolCirc}^{\dagger}$ over $\mathbf{FRel}_2$, that is as Boolean relations rather than functions. In this case B again defines a valid interpretation $\mathbf{BoolCirc}^{\dagger} \to \mathbf{FRel}_2$. The resulting theory is a model of non-deterministic computation.

In any theory, a process of type $p : I \to A$ is called a *point*, or sometimes a *state*, of A. Dually, a process of type $e : A \to I$ is called a *co-point*, or sometimes an *effect* on A. For example, in **FHilb** the points $\psi : I \to A$ are in one-to-one correspondence with the vectors of A, while in **FRel** a point $s : I \to X$ is precisely a subset of X.

In a theory with a dagger the set of points is isomorphic to the set of copoints (or in other language, for every state there is a corresponding effect and vice versa). This allows us to define another important concept.

Definition 2. Given two points $\psi, \phi : I \to A$ we define their *inner product* as $\phi^{\dagger} \circ \psi$. Dually, the outer product is defined as $\phi \circ \psi^{\dagger}$.

As one may expect, the inner product is always a scalar. The diagrammatic language automatically allows the same tricks—and more— as Dirac notation does in Hilbert spaces. Indeed one can view the diagrammatic language as a 2-dimensional generalisation of Dirac notation.

Example 11 (Finite-dimensional Hilbert spaces). In **FHilb** the inner product defined by the dagger, is exactly the usual inner product $\langle \phi | \psi \rangle$.

Example 12 (Sets and Relations). In **FRel** the inner product $r^{\dagger} \circ s$ is 0 if the r and s are disjoint as subsets, and 1 otherwise.

2 Pure state quantum mechanics

2.1 The elements of an operational theory

It is remarkable that the the basic language of quantum mechanics—states, effects, unitarity, self-adjointness, inner products, tensor products—can all be defined in the abstract setting of generalised compositional theories. We now have enough material to describe a formal operational framework for pure state quantum mechanics in purely diagrammatic terms.

- A *preparation* is any process which produces a state; that is to say it is process of type $p : I \to A$.

Preparations are not restricted to producing single systems; a preparation process of type $I \to A_1 \otimes \cdots A_n$ is called *multipartite*. Of course, multipartite preparations need not be separable.

When interpreted in **FHilb** each preparation process yields a ray in some Hilbert space, which, ignoring global phase, we may identify with a specific quantum state. It may happen, depending on the equations of the formal theory, that different preparation processes produce the same state.

- A *transformation* is any process which acts on states and produces new states, and which is unitary:

Once again, transformations may act on one or many systems at the same time.
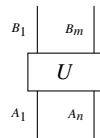
- Measurements are processes which accept quantum inputs and produce classical information about the state which was input. Since, for now, our theory only has pure states, we will work with *non-degenerate post-selected* measurements[6]; i.e. we know that a definite outcome has occurred, and that outcome corresponds to a definite quantum state. Therefore, measurements are one-dimensional effects, represented as co-points:

[6] In other words, rank 1 projectors.
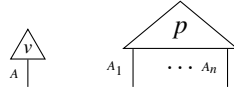
The classical information is implicit in the choice of copoint, and hence not represented. Since copoints do not have quantum outputs, these processes correspond to *demolition measurements*, where the original system is consumed by the measurement process. However, by combining an effect with the corresponding state preparation we can also represent non-demolition measurements:
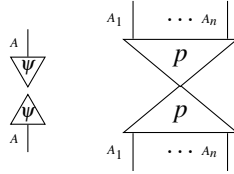
To properly represent the non-determinism of quantum measurements we need to consider mixed states; this is dealt with in Section 4. More general measurements can be represented within the theory, however they will not be described here.

This basic recipe—preparations, transformations, and measurements—allows any experimental setup to be described in terms of the processes which realise it. More precisely, since we use post-selected measurements, the diagram really represents a *run* of the experiment where a certain outcome occurred. We call an experiment *closed* when it has no external inputs or outputs. Any closed experiment is necessarily described by a process of type $x : I \rightarrow I$; that is, a scalar. This scalar is the abstract counterpart to the probability amplitude for performing the process and observing the specified result. Indeed, when such a diagram is interpreted in **FHilb**, the result is exactly the probability amplitude.

Example 13. The theory **QuCirc** has the structure described above, and we can use it to define a simple experiment. For example, the diagram below corresponds to preparing a qubit in the $|0\rangle$ state, applying a unitary gate to it, and, upon measuring in the computational basis, finding that the qubit is in the state $|1\rangle$.
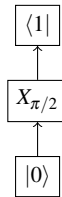
Using the interpretation map $[\![\cdot]\!] : \mathbf{QuCirc} \rightarrow \mathbf{Qubit}$ we can calculate the amplitude for this experimental result.

$$\langle 1| \circ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \circ |0\rangle \quad = \quad \frac{-i}{\sqrt{2}}$$

To summarise the elements of the framework, a formal generalised compositional theory consists of:

- A collection of basic systems and processes, corresponding to the available "lab equipment".
- The collection of all diagrams constructed from the basic processes, corresponding to every possible experiment that could be built from the given equipment. We consider diagrams modulo topological equivalence: equivalent diagrams correspond to the same experiment.
- A (possibly empty) collection of axioms, presented as rewrite rules over diagrams, which specify behavioural equivalence of processes. These rules tell us when a piece of the experimental setup can safely be replaced by another without changing the result of the experiment.
- Finally, given the above, we'll usually consider (sound) *interpretation maps* of the formal theory into some concrete mathematical structure, such as Hilbert spaces.

So far we have been operating at an extremely high level of generality. To focus our attention on quantum systems we will now gradually introduce more structure to our theories. We identify certain structural features of the Hilbert space presentation of quantum mechanics, and provide an abstract realisation of those features in terms of basic processes and equations, whose behaviour reproduces various quantum phenomena in the abstract setting of generalised compositional theories.

The rest of this section will layout which basic processes and equations we will need to realise. As we do so, we'll say goodbye to some of the models introduced earlier, but the two most important ones, **FHilb** and **FRel**, will still be applicable.

2.2 Duals

The next piece of structure that will be required is the existence of *duals*[7].

Definition 3. A system A has a *dual* if there exists a system A^* and processes

$$e_A : I \to A^* \otimes A \quad \text{and} \quad d_A : A \otimes A^* \to I$$

such that we have the following equations:

---

[7] For the experts in category theory, this additional structure can be summed up by saying we operate in a dagger-compact category, rather than just a symmetric monoidal category.

$$(d_A \otimes 1_A) \circ (1_A \otimes e_A) = 1_A \qquad (1_{A^*} \otimes d_A) \circ (e_A \otimes 1_{A^*}) = 1_{A^*}$$

Since this definition is rather hard to parse we will immediately move to its diagrammatic form. We indicate the dual system A^* by a wire labelled by A but directed in the opposite direction. The maps d_A and e_A are represented by wires with half turns, henceforth "caps" and "cups". The equations above then take the form of "straightening wires":

$$e_A := \quad d_A :=$$

In general a system might have more than one dual, but they are all guaranteed to be isomorphic. We'll assume that every system has a given dual, and in particular $(A \otimes B)^* = B^* \otimes A^*$, in which case $d_{A \otimes B}$ and $e_{A \otimes B}$ take the form of nested caps and cups. Furthermore, we'll assume that the double dual $A^{**} = A$. These simplifications automatically hold in any theory presented diagrammatically; taking them as the general case saves a lot of bureaucracy.

Example 14 (Finite dimensional Hilbert spaces). Let A be a Hilbert space of dimension d, then A^* is the usual dual space; that is, the space of linear functionals from A to the complex numbers. Supposing that $\{|a_i\rangle\}$ is a basis for the space A, then the cup and cap are given by the linear maps

$$e_A : 1 \mapsto \sum_i \langle a_i| \otimes |a_i\rangle \ , \qquad d_A : \sum_i |a_i\rangle \otimes \langle a_i| \mapsto 1 \ .$$

However since we are in a finite-dimensional setting we could also choose $A^* = A$; specialising to the case of qubits, we can now view the cup and cap as the preparation and projection onto a Bell state:

$$e_Q = |00\rangle + |11\rangle \qquad d_Q = \langle 00| + \langle 11|$$

Recall the quantum teleportations protocol: Alice has some unknown state that she wishes to send to Bob, but they do not share a quantum channel. However they have a classical channel, and have previously shared a Bell pair. In order to send her qubit to Bob, Alice measures her two qubits in the Bell basis, and transmits the result to Bob. Now Bob simply applies some unitary map (depending on Alice's outcome) to his half of the Bell pair to recover the qubit that Alice wanted to send. Since, for the moment, we are operating in a post-selected setting, we'll assume that Alice observes the outcome corresponding to the state $\Phi_+ = (|00\rangle + |11\rangle)/\sqrt{2}$ at her measurement. In this case Bob need do nothing to his qubit. The whole set up is shown below:
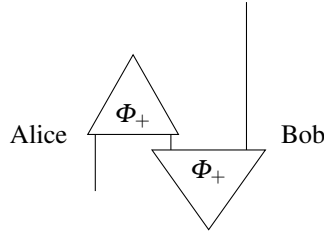
Knowing that the projection onto Φ_+ is just the effect d_Q, we can rewrite the protocol as shown and demonstrate the protocol purely diagrammatically:
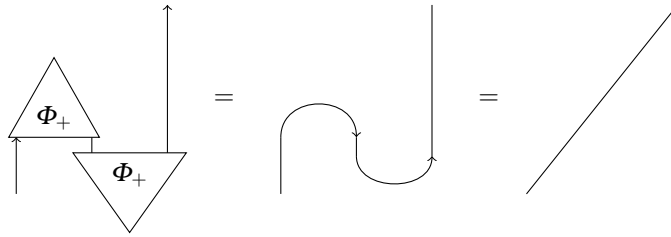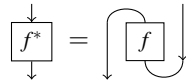

Example 15 (Sets and Relations). In **FRel**, the dual of a set X is just the same set X again. The cup is given by the "name" of the identity:

$$e_X = \{(*, (x,x)) \mid x \in X\}$$

while the cap, d_X is just the converse of e_X.

Using caps and cups, we can turn any process $f : A \to B$ into a process on the dual objects going in the opposite direction: $f^* : B^* \to A^*$.

This is sometimes called the transpose of f, but this terminology can be misleading. In **FHilb**, f^* is the map that takes a linear form $\langle\xi| \in B^*$ to $\langle\xi|f \in A^*$. We refer to this map simply as the *upper-star* of f. Clearly, we have $f^{**} = f$.

It is also required that the dagger and the duals interact nicely. More precisely we have the equations:

In any theory with both a dagger and duals, we can define a third operation, the *lower-star* of f as $f_* := (f^\dagger)^* = (f^*)^\dagger$. Again this is involutive, i.e. $f_{**} = f$. We'll return to the uses of the upper and lower stars in Section 4.

Finally, the cup and cap can be used to define a *trace* in purely diagrammatic terms:

$$\mathrm{Tr}(\boxed{\ f\ }) = \boxed{\ f\ }$$

Checking the Hilbert space interpretation, it is easy to see that this coincides with the usual definition.

$$\mathrm{Tr}(\boxed{\ f\ }) = (\sum_i \langle ii|) \circ (1_A \otimes f) \circ (\sum_j |jj\rangle) = \sum_i \langle i|f|i\rangle = \sum_i f_{ii}$$

In the diagrammatic form it is trivial to prove that trace is invariant under cyclic permutation:

$$\mathrm{Tr}\left(\boxed{\begin{matrix} g \\ f \end{matrix}}\right) = \boxed{\begin{matrix} g \\ f \end{matrix}} = \boxed{g}\ \boxed{f} = \boxed{\begin{matrix} f \\ g \end{matrix}} = \mathrm{Tr}\left(\boxed{\begin{matrix} f \\ g \end{matrix}}\right)$$

The partial trace can be defined analogously.

$$\mathrm{Tr}_B^A\left(\boxed{\ U\ }\right) = \boxed{\ U\ }$$

By adding duals we have enlarged the class of possible diagrams, since wires may now loop back from inputs to outputs and vice versa, but the basic principle of diagram equality does not change: **Two diagrams are considered equal if one can be smoothly transformed to another, by bending, stretching, or crossing wires, and moving boxes around**. With this in mind we can update the key theorem.

Theorem 2 (Fundamental Theorem of Diagrams with Daggers and Duals). *Given any two generalised compositional theories* **C** *and* **D** *with daggers and duals, and a map* $F : \mathbf{C} \to \mathbf{D}$, *for any two diagrams* d *and* d' *in* **C**, *if* $d = d'$ *as diagrams then* $Fd = Fd'$ *in* **D**.

Once again, the full details are found in [19].

Remark 5. We need not demand any additional conditions on the class of mappings to guarantee the preservation of duals; since they are defined in terms of processes, the structure is automatically preserved.

2.3 Observable Structures

An *observable* yields classical data from a physical system [22]. The key distinction between classical and quantum data is that classical data may be freely copied and deleted, while this is impossible for quantum data, due to the no-cloning [23, 24] and no-deleting [25] theorems.

In quantum mechanics, an observable is represented by a self-adjoint operator. This (non-degenerate) operator encodes certain classical data as its orthonormal basis of eigenstates, the possible outcomes of the corresponding measurement. Note that if a quantum state is known to be a member of a given orthonormal basis, such as the eigenbasis $\{|a_i\rangle\}$ of some observable, then it *can* be copied and deleted via the maps

$$\delta : |a_i\rangle \mapsto |a_i\rangle \otimes |a_i\rangle \quad \text{and} \quad \varepsilon : |a_i\rangle \mapsto 1.$$

Hence we can view the classical content of a quantum measurement as the possibility to copy and delete its set of outcomes. We will axiomatise quantum observables by describing the copying and deleting operations as algebraic structures inside a general compositional theory. The relevant structure is called a †-*special commutative Frobenius algebra*, and we will now build up its definition one piece at a time.

Definition 4. A *commutative monoid* in **C** is a triple (X, μ, η), where μ and η are maps

$$\mu : X \otimes X \to X \qquad\qquad \eta : I \to X$$

which we write graphically as $\mu = \text{⨁}$, $\eta = \text{○}$. These operations satisfy the following equations:
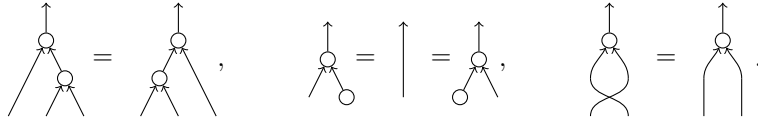

Remark 6. The process μ can be understood as a *multiplication* for systems of type X; the first and last equations assert that this operation is associative and commutative respectively. The process η is the *unit* for this multiplication: the second equation asserts that multiplication by the unit is simply the identity.
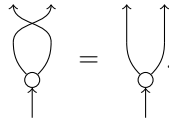
The dual to a monoid is a *comonoid*.

Definition 5. A *comonoid* in a theory **C** consists of a triple (X, δ, ε) where δ and ε are processes

$$\delta : X \to X \otimes X \qquad\qquad \varepsilon : X \to I$$

satisfying the equations of Definition 4 but in reverse, viz:
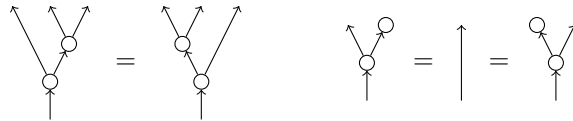
A comonoid is *cocommutative* if it satisfies:

The processes δ and ε are called the *comultiplication* and *counit* respectively.

Example 16. We have already met the basic example of a comonoid: in **FHilb**, for any orthonormal basis $\{x_i\}_i$ of a space X we obtain a comonoid via 'copying' and 'erasing' processes mentioned above:

$$\delta : x_i \mapsto x_i \otimes x_i \qquad \varepsilon : x_i \mapsto 1$$

Remark 7. Thanks to the dagger, if (X, δ, ε) is a comonoid then $(X, \delta^\dagger, \varepsilon^\dagger)$ is automatically a monoid, and vice versa.

Generally speaking, a process is called a *homomorphism* if it preserves some algebraic structure. In the context of GCTs, such preservation is usually expressed by a process commuting with another which reifies that structure. For example:

Definition 6. Given two comonoids (X, δ, ε) and $(X', \delta', \varepsilon')$, a *comonoid homomorphism* is a process $f : X \to X'$ such that

$$\delta' \circ f = (f \otimes f) \circ \delta \quad \text{and} \quad \varepsilon' \circ f = \varepsilon .$$
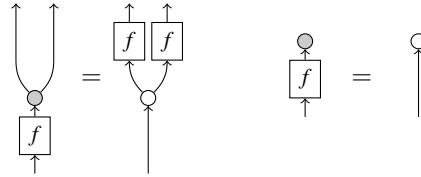

Monoid homomorphisms are defined similarly.

Remark 8. The definition above is the most general, but we will frequently encounter cases where $f : X \to X$ is homomorphism between two comonoids defined on the same object, or from a single comonoid to itself.

The structures of greatest interest for this paper are algebras containing both monoids and comonoids.

Definition 7. A *commutative Frobenius algebra* is a 5-tuple $(X, \delta, \varepsilon, \mu, \eta)$ where
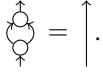
1. (X, δ, ε) is a cocommutative comonoid;
2. (X, μ, η) is a commutative monoid; and,
3. δ and μ satisfy the following equations:
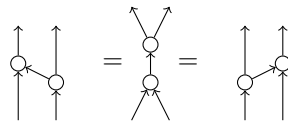
Finally, we can define:

Definition 8. A †-*special Frobenius algebra* (†-SCFA) is a commutative Frobenius algebra

$$\mathscr{O}_\circ = (\mu_\circ : X \otimes X \to X, \ \eta_\circ : I \to X,$$
$$\delta_\circ : X \to X \otimes X, \ \varepsilon_\circ : X \to I)$$

such that $\delta_\circ = (\mu_\circ)^\dagger$, $\varepsilon_\circ = (\eta_\circ)^\dagger$ and $\Large\substack{\uparrow \\ \otimes \\ \otimes \\ \uparrow} = \Big\uparrow$.

The preceding definitions may seem rather opaque, and not fully justified by the intuition that a quantum observable is somehow encoded by the maps which copy and delete its eigenstates. However complex it may appear (and we shall shortly simplify it), the importance of Definition 8 rests on the fact [26] that in **FHilb** *every* †-SCFA arises from a comonoid defined by copying an orthonormal basis as described above. Since orthonormal bases define non-degenerate quantum observables, †-SCFAs are also called *observable structures*.

Concretely, given an orthonormal basis $\{|i\rangle\}_i$ then $\delta_\circ :: |i\rangle \mapsto |ii\rangle$ defines an observable, and all observables are of this form for some orthonormal basis. The resulting intuition is that $\delta_\circ$ is an operation that 'copies' basis vectors, and that $\varepsilon_\circ$ 'erases' them [22]. We will use the symbolic representation $(\mu_\circ, \eta_\circ, \delta_\circ, \varepsilon_\circ)$ and the pictorial one $(\substack{\uparrow \\ \otimes}, \dot{\circ}, \substack{\otimes \\ \downarrow}, \varphi)$ interchangeably.

Example 17 (Sets and Relations). Perhaps surprisingly, **FRel** also has many distinct observable structures, which have been classified by Pavlovic [27]. Even on the two element set there are two, namely

$$\delta_\circ : i \mapsto \{(i,i)\}$$
$$\delta_\circledcirc : \begin{cases} 0 \mapsto \{(0,0),(1,1)\} \\ 1 \mapsto \{(0,1),(1,0)\} \end{cases}$$

In fact, this pair is *strongly complementary* in the sense to be explained in Section 3.

Manipulating observable structures in the graphical language is extremely convenient since they obey a remarkable normal form law. Let $\delta_n : X \to X^{\otimes n}$ be defined by the recursion
$$\delta_0 := \varepsilon \qquad \delta_{n+1} := (\delta_n \otimes 1_A) \circ \delta$$

and define μ_m analogously. Now we have the following important theorem:

Theorem 3. *Given an SCFA* $(X, \delta, \varepsilon, \mu, \eta)$ *let* $f : X^{\otimes m} \to X^{\otimes n}$ *be a map constructed from* δ, ε, μ *and* η *whose graphical form is connected. Then* $f = \delta_n \circ \mu_m$.
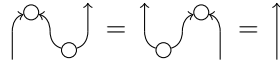
Proposition 1. Given an observable structure $\mathscr{O}_\circ$ on X, let $(\bigcirc)_n^m$ denote the '(n,m)-legged spider':
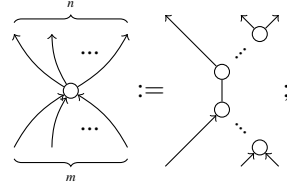
then any morphism $X^{\otimes n} \to X^{\otimes m}$ built from $\mu_\circ, \eta_\circ, \delta_\circ$ and $\varepsilon_\circ$ via †-SMC structure which has a connected graph is equal to the $(\circ)_n^m$. Hence, spiders compose as follows:

$$
\tag{2}
$$
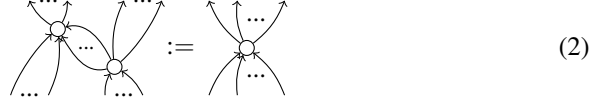
Thanks to the spider rule (2), every observable structure on X makes X dual to itself (in the sense of Definition 3), via the cup and cap:

The upper-star with respect to this cup and cap corresponds in **FHilb** to transposition in the given basis. For that reason, we call this the $\circ$-transpose $f^{\circ T}$. The lower star corresponds to complex conjugation in the basis of $\mathscr{O}_\circ$, so we call it the $\circ$-conjugate $f_\circledast := (f^{\circ T})^\dagger$.

Recall that a process $k : I \to X$ is called a *point of X*. In **FHilb** the points of X are simply vectors in the Hilbert space X. The abstract analogue of the eigenvectors of an observable in **FHilb** are the *classical points* of an observable structure.

Definition 9. A *classical point* for an observable structure is a state that is copied by the comultiplication and deleted by the counit:

$$
\tag{3}
$$
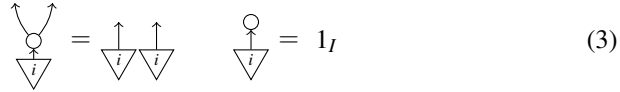
We will depict classical points as triangles of the same colour as their observable structure.

Remark 9. Another way to say the same thing, is to define classical points as comonoid homomorphisms from the trivial comonoid $(I, 1_I, 1_I)$ to (X, δ, ε).

In quantum computing, it is common to think of elements of a product basis as strings of some kind. E.g. for qubits:

$$|010011\rangle \leftrightarrow |0\rangle \otimes |1\rangle \otimes |0\rangle \otimes |0\rangle \otimes |1\rangle \otimes |1\rangle$$

Such product bases are precisely the classical points of *products of observable structures*. Given an observable structure $\mathscr{O}_\circ$ on X, and another $\mathscr{O}_\bullet$ on Y, we form a new observable structure on $X \otimes Y$ by taking their product:

$$\delta = \qquad\qquad \varepsilon =$$

Evidently any pair of classical points for the constituent observable structures will be copied.

Generalising, the classical points of any n-ary product of observable structures are nothing more than lists of classical points, one for each factor.

Working concretely in Hilbert space, one can use the linear structure to give another set of equations for observable structures. Consider some basis vector $|i\rangle$, then the map $|ii\rangle\langle i|$ has the diagrammatic form:

But notice that the sum $\sum_i |ii\rangle\langle i|$ is nothing more than the the map $\delta : |i\rangle \to |ii\rangle$. A similar statement can be made for the counit ε. Hence given the complete set of classical points for an observable structure $\mathscr{O}_\circ$ we have the following equations:

Indeed these can be generalised to arbitrary spiders:

Note that generalised compositional theories do not necessarily admit addition of diagrams: we introduce these equations as way of generalising from concepts defined in Hilbert space to the abstract setting where there need not be any linear structure.

Linear maps have the property that if two maps are equal on every element of a basis, the maps themselves are equal. In analogy to this we define the following:

Definition 10. Let $\mathscr{O}_\circ$ be an observable structure on X, with classical points $\{k_i\}_i$; we say that $\mathscr{O}_\circ$ has *enough classical points* if, for every system Y, and every pair of processes $f, g : X \to Y$, we have

$$(\forall i : f \circ k_i = g \circ k_i) \Rightarrow f = g \ .$$

This property does not necessarily hold in an arbitrary GCT (although obviously it does in **FHilb**) however when it does hold certain statements can be made stronger. For example, many implications described in the subsequent sections are equivalences if the underlying object has enough classical points.

2.4 Phase Group for an Observable Structure

Let ψ and ϕ be two points of X. Given an observable structure $\mathscr{O}_\circ$ on X, applying the multiplication $\mu_\circ$ to ψ and ϕ yields another point of X:
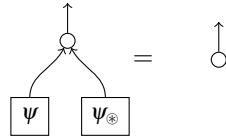
$$\psi +_\circ \phi \ := \ \mu_\circ(\psi \otimes \phi)$$
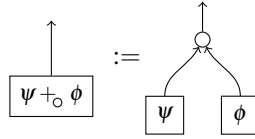
Since $\mu_\circ$ is commutative and associative, and it has a unit point (namely $\eta_\circ$), the operation $+_\circ$ gives the points of X the structure of a commutative monoid.

If we restrict to those points $\psi : I \to X$ which satisfy

$$\psi +_\circ \psi_\circledast = \eta_\circ$$

we obtain an abelian group $\Phi_\circ$, called the *phase group* of $\mathscr{O}_\circ$ [28, 29]. The elements of this group are called *phases*. Note the phase group is non-empty, since the unit $\eta_\circ$ satisfies the required equation. We let $-\alpha := \alpha_\circledast$ for phases, and represent these points as circles with one output, labelled by a phase.

Example 18. In **FHilb**, let $\mathscr{O}_\circ$ be defined by some orthonormal basis $\{|i\rangle\}_i$. One can verify by direct calculation that a vector $|\psi\rangle$ lies in the phase group $\Phi_\circ$ if and only if we have $|\langle i|\psi\rangle|^2 = 1/D$, for all i, where D is the dimension of the underlying space. Such vectors are called *unbiased* for the basis $\{|i\rangle\}_i$. The multiplication is then the convolution (pointwise) product.

Concretely, for a qubit observable given by $\mu_\circ = |0\rangle\langle 00| + |1\rangle\langle 11|$, the phases are the unbiased states, which are all of the form:

$$|\alpha\rangle = \begin{pmatrix} 1 \\ e^{i\alpha} \end{pmatrix} ,$$

with the multiplication:

$$\mu_\circ \left(\begin{pmatrix} 1 \\ e^{i\alpha} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ e^{i\beta} \end{pmatrix} \right) = \begin{pmatrix} 1 \\ e^{i(\alpha+\beta)} \end{pmatrix} .$$

We therefore see that the phase group for a qubit observable is the circle group. It is an easy exercise to check that for a D-dimensional Hilbert space the phase group for any observable is isomorphic to the $(D-1)$-dimensional torus.

The name 'phase group' comes from fact that the elements of the $\Phi_\circ$ correspond to unitary maps which preserve the basis defining $\mathcal{O}_\circ$. We can map any point $\psi : I \to X$ onto an operation on X via the *left action*, $\Lambda_\circ(\psi) = \mu_\circ \circ (\psi \otimes 1_X)$, or in pictures:
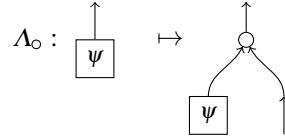
Then we have the following facts:

Proposition 2. *Let $\phi, \psi \in \Phi_\circ$; then*

1. $\Lambda_\circ(\phi)$ is unitary;
2. $\Lambda_\circ(\phi) \circ \Lambda_\circ(\psi) = \Lambda_\circ(\phi + \psi) = \Lambda_\circ(\psi) \circ \Lambda_\circ(\phi)$
3. If k is a classical point for $\mathcal{O}_\circ$ then $\Lambda_\circ(\phi) \circ k = k \otimes s$ for some scalar s.

Proof. 1. We show that $\Lambda_\circ(\psi)^\dagger \circ \Lambda_\circ(\psi) = 1$:
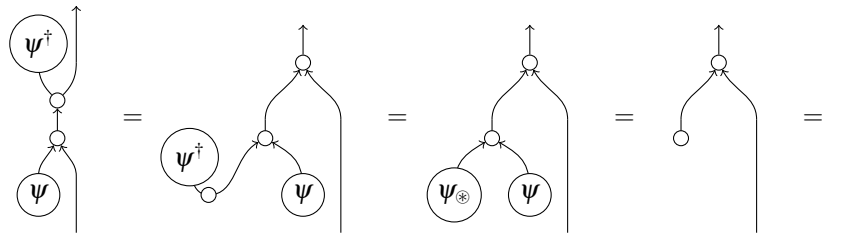
The first equation is the spider rule while the second is the definition of $\psi_\circledast$. The case $\Lambda_\circ(\psi) \circ (\Lambda_\circ(\psi))^\dagger = 1$ is similar.
2. This follows immediately from the associativity and commutativity of $\mu_\circ$:
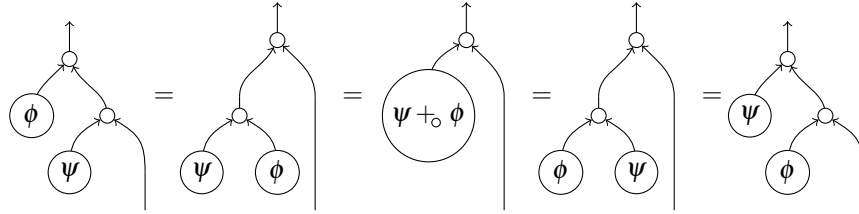
3. This follow from the definition of classical points.
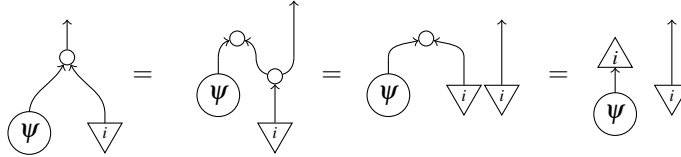
The image $\Lambda_\circ(\Phi_\circ)$ is therefore an abelian subgroup of the unitaries on X, which is isomorphic to $\Phi_\circ$. We refer to these as *phase maps*. If we reinterpret the third part of the preceding proposition in terms of linear algebra, we see that every classical point of $\mathscr{O}_\circ$ is an eigenvector of every phase map in $\Lambda_\circ(\Phi_\circ)$. This in turn "explains" why they commute with each other.

Example 19. Let $\mathscr{O}_\circ$ be defined by $\mu_\circ = |0\rangle\langle00| + |1\rangle\langle11|$ as above. Now for $\alpha \in \Phi_\circ$ we have

$$\Lambda_\circ(\alpha) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}$$

Hence the phase group in Hilbert spaces is exactly the group of phase shifts relative to the given basis.

Generalising from the preceding discussion, we can now introduce 'spiders decorated with phases':

(4)

which compose as follows:

(5)

In the following sections we will refer to this generalised composition rule for phased spiders as *the spider law*.

2.5 Two toy models

In this section we'll introduce two "toy" models of quantum mechanics. The first is the restriction of quantum mechanics to stabilizer states; this theory we call **Stab**. The second is the toy model due to Spekkens [30], which we refer to as **Spek**. While the first of these is indeed a true subtheory of quantum mechanics, **Spek** is a local hidden variable model. By casting both of these in the language of generalised compositional theories we can see that the difference between is in fact very slight.

Before discussing these concrete models, we'll introduce a formal precursor theory. Let **Toy** be the general compositional theory built from the formal generators:

- one basic system, which we denote T;
- six points $z_0, z_1, x_0, x_1, y_0, y_1 : I \to T$ and their corresponding copoints;
- 24 unitary maps $T \to T$ which form a group isomorphic to the symmetric group S_4;
- one observable structure $\mathcal{O}_\circ$, whose classical points are z_0 and z_1, and whose phase group comprises the remaining four points.

Note that **Toy** is not fully specified: to do so we ought to say which group the phase group is, and how the corresponding unitaries embed into the endomorphisms of T. Since $\Phi_\circ$ is a four-element group we have only two choices here: $\mathbb{Z}_4$, or $\mathbb{Z}_2 \times \mathbb{Z}_2$. As we will see this choice will make the difference between stabilizer quantum mechanics and the quantum-like local hidden variable theory.

Let **Stab** be the subtheory of **FHilb** generated by the following elements:

- One basic system $\mathbb{C}^2$, which we call Q.
- Six points $I \to Q$:

$$z_0 = |0\rangle \qquad x_0 = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad y_0 = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$$

$$z_1 = |1\rangle \qquad x_1 = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \qquad y_1 = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$$

- The group of unitaries generated by the matrices:

$$Z_{\pi/2} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \qquad X_{\pi/2} = \frac{1}{\sqrt{-2i}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$$

  This group is known in the quantum computation literature as the Clifford group for one qubit; it is isomorphic to S_4. The other key property of this group is that it acts as a permutation on the states defined above, so we cannot generate new states via unitaries.
- An observable structure $\mathcal{O}_\circ$ defined by the basis $|0\rangle, |1\rangle$.

Evidently the classical points of $\mathcal{O}_\circ$ are indeed z_0 and z_1 and the remaining points are unbiased for this basis, hence part of $\Phi_\circ$. One can check that

$$\Lambda_\circ(y_0) = Z_{\pi/2}$$

and which generates a four element cyclic subgroup, hence the phase group $\Phi_\circ$ is $\mathbb{Z}_4$.

We now introduce Spekkens' toy theory. The toy theory is a local hidden variable theory, based on epistemic restrictions. There is a single basic system, the toy bit, which can have one of four possible states. We formalise the state space simply as a four-element set. However, we now impose the epistemic restriction that any state preparation (and, dually, measurement) may only narrow down the state to two of the possible four. Hence the "states" of the toy theory are two-element subsets. Although Spekkens' original presentation [30] was informal, the toy theory is ideally studied as subtheory of **FRel**. Following [31], let **Spek** be the subtheory of **FRel** generated by the following elements:

- One basic system, the four element set $\mathbf{4} = \{0,1,2,3\}$.
- Six points:

$$
\begin{array}{lll}
z_0 = \{0,1\} & x_0 = \{0,2\} & y_0 = \{0,3\} \\
z_1 = \{2,3\} & x_1 = \{1,3\} & y_1 = \{1,2\}
\end{array}
$$

- The full group of permutations on $\mathbf{4}$;
- An observable structure $\mathscr{O}_\circ$ defined by

$$
\mu_\circ : \begin{array}{l}
\{00,11\} \sim 0 \\
\{01,10\} \sim 1 \\
\{22,33\} \sim 2 \\
\{23,32\} \sim 3
\end{array}
\qquad \eta_\circ : * \sim \{0,2\}
$$

where we write the tensor as juxtaposition, i.e. $00 = (0,0)$.

Once again we easily check that the classical points for $\mathscr{O}_\circ$ are z_0 and z_1, and the other four form the phase group $\Phi_\circ$. The phase group in this case is generated by the transpositions $(0\ 1)$ and $(2\ 3)$; hence $\Phi_\circ \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

As should be evident by this point both **Stab** and **Spek** are realisations of the incomplete theory **Toy**. The only notable difference between them is the group structure of $\Phi_\circ$. This highlights the importance of the phase group for understanding non-locality in generalised compositional theories.

Remark 10. In the description above the group of unitaries was given *a priori*. This is not necessary. If we include a second observable structure $\mathscr{O}_\circ$, corresponding to the classical points x_0 and x_1, the the union of the two phase groups $\Phi_\circ$ and $\Phi_\circ$ yields *all* unitaries described above. These two observables are complementary in the sense described below. Hence these two theories are in a sense minimal.

3 Complementarity and Strong Complementarity

In the Hilbert space presentation of quantum mechanics, two observables are *complementary* if their bases of eigenstates are mutually unbiased. That is, for any i, j, $|\langle v_i | v'_j \rangle|^2 = 1/D$. In the graphical notation:

$$\text{[diagram]} = \tfrac{1}{D}$$

A question posed by Coecke and Duncan [28, 29] was, "Can we represent complementarity purely in terms of interacting observable structures?" It turns out that complementarity is equivalent to a simple diagrammatic equation. First, we can move $1/D$ in the above equation to the other side and express it as a circle, as the trace of the identity always equals D. Then, replace 1 on the RHS with "deleted points".

$$\text{[diagram]} = \text{[diagram]} \tag{6}$$

As we saw in section 2.3, observable structures fix an isomorphism of a space with its dual space, via the transpose. While it is not true in general that $|\psi\rangle^{\bigcirc\mathrm{T}} = \langle\psi|$, the transpose does take classical points for a particular observable structure to their adjoints:

$$|v_i\rangle^{\bigcirc\mathrm{T}} = \langle v_i | \quad\text{and}\quad |v'_j\rangle^{\bigcirc\mathrm{T}} = \langle v'_j | \ .$$

Graphically:

$$\text{[diagram]} = \text{[diagram]} \qquad \text{[diagram]} = \text{[diagram]} \tag{7}$$

Exercise. Prove this.

We can rewrite the left hand side of equation (6) using this fact.

$$\text{[diagram]} = \text{[diagram]} = \text{[diagram]} = \text{[diagram]} \tag{8}$$

The last equation follows by substituting the symbol S for its definition, viz:

$$\text{[diagram]} \; S \; = \; \text{[diagram]} \tag{9}$$

Unifying Equations (6) and (8) we have:

$$ \tag{10} $$

Since this equation holds for all classical points i and j, if we now appeal to the fact that **FHilb** has enough classical points (cf. Definition 10), we can conclude that identity holds without points:

$$ \tag{11} $$

Remark 11. The above equation is (up to a scalar factor) one of the defining equations a *Hopf algebra*, in which case the map *S* is called the *antipode*. For that reason, we refer to (11) as the *Hopf law*. As we will see in the next section, subject to some additional assumptions, pairs of complementarity observables do indeed form Hopf algebras with the antipode defined as in Equation (9).

Notice if we assume Equation (11) we can derive Equation (6) without any additional assumptions. In other words, if $\mathcal{O}_{\circ}$ and $\mathcal{O}_{\odot}$ satisfy the Hopf law their classical points are mutually unbiased. Thus, we take the Hopf law to be the defining equation for our *abstract* notion of complementarity.

Definition 11. A pair $(\mathcal{O}_{\circ}, \mathcal{O}_{\odot})$ of observables on the same object X is *complementary* iff:

$$ \qquad = \qquad \text{where} \qquad \widehat{S} = \quad . $$

Since every observable in **FHilb** has enough classical points, Definition 11 is equivalent[8] to saying that observables are complementary if their eigenbases are mutually unbiased with respect to the other. (See [28] for more details). Hence, we reclaim the usual notion of quantum complementarity, and extend it to a more general setting.

Definition 12. A pair $(\mathcal{O}_{\circ}, \mathcal{O}_{\odot})$ of observables on the same object X is *coherent* iff:

$$ \qquad = \qquad\qquad \qquad = \qquad\qquad = \quad . $$

---

[8] Indeed Equations (6) and (11) are equivalent in any theory wherever at least one of the observable structures has enough classical points.

In other words, the unit point $\eta_\circ$ ($\scriptstyle\bigcirc\!\!\!\!\;$) is, modulo a scalar factor, a classical point for $\mathscr{O}_\circ$, and vice versa.

$$\exists i,j : \quad \begin{smallmatrix}\circ\\\circ\end{smallmatrix} \;\; \underset{i}{\bigvee}\!\!\!\!\uparrow \; = \; \uparrow\!\!\!\circ \qquad\qquad \begin{smallmatrix}\circ\\\circ\end{smallmatrix}\;\; \underset{j}{\blacktriangledown}\!\!\!\!\downarrow \; = \; \uparrow\!\!\!\circ$$

We will assume that the scalar $\begin{smallmatrix}\circ\\\circ\end{smallmatrix}$ is always cancellable.

Example 20. Consider the two observables on the Hilbert space $\mathbb{C}^2$ corresponding to the Z and X spins:

$$\delta_\circ : \begin{aligned}|0\rangle &\mapsto |00\rangle\\ |1\rangle &\mapsto |11\rangle\end{aligned} \qquad\qquad \delta_\bullet : \begin{aligned}|+\rangle &\mapsto |++\rangle\\ |-\rangle &\mapsto |--\rangle\end{aligned}$$

$$\eta_\circ : \begin{aligned}|0\rangle &\mapsto 1\\ |1\rangle &\mapsto 1\end{aligned} \qquad\qquad \eta_\bullet : \begin{aligned}|+\rangle &\mapsto 1\\ |-\rangle &\mapsto 1\end{aligned}$$

Computing $\eta_\circ$ we obtain:

$$\eta_\circ = (\varepsilon_\circ)^\dagger = \left[1 \mapsto (|0\rangle + |1\rangle)\right] = \sqrt{2}\,|+\rangle$$

which is indeed proportional to a classical point for $\delta_\bullet$. By a similar computation we obtain $\eta_\bullet = \sqrt{2}\,|0\rangle$, from which the value of their inner product $\begin{smallmatrix}\circ\\\circ\end{smallmatrix} = \sqrt{2}$ follows. The equations of definition 12 can easily be verified from here, demonstrating that $\mathscr{O}_\circ$ and $\mathscr{O}_\bullet$ are coherent.

Proposition 3. In **FHilb** if $O_\circ$ and $O_\bullet$ are self-adjoint operators corresponding to complementary observables, one can always choose a pair of coherent observable structures $(\mathscr{O}_\circ, \mathscr{O}_\bullet)$ whose classical points correspond to the eigenbases of $O_\circ$ and $O_\bullet$.

Proof. Let $\{|a_i\rangle\}_{i=1}^n$ and $\{|b_j\rangle\}_{j=1}^n$ be orthonormal eigenbases for $O_\circ$ and $O_\bullet$ respectively. Since the bases are mutually unbiased we have

$$|b_j\rangle = \frac{1}{\sqrt{n}}\left[\alpha_{1j}|a_1\rangle + \cdots \alpha_{nj}|a_n\rangle\right]$$

where the α_{ij} are scalars satisfying $|\alpha_{ij}| = 1$. Setting $|a_i'\rangle = \alpha_{i1}|a_i\rangle$, we see that $\{|a_i'\rangle\}_i$ is also an orthonormal eigenbasis for $O_\circ$, which is still mutually unbiased with respect to $\{|b_j\rangle\}_j$. Now define:

$$\delta_\circ : |a_i'\rangle \mapsto |a_i'\rangle \otimes |a_i'\rangle$$
$$\varepsilon_\circ : |a_i'\rangle \mapsto 1$$

This choice yields $\eta_\circ = (\varepsilon_\circ)^\dagger = \sum_i |a_i'\rangle = \sqrt{n}\,|b_1\rangle$.

In a similar fashion we can choose an eigenbasis $|b'_1\rangle, \ldots, |b'_n\rangle$ for $O_\circledcirc$ such that the resulting $\delta_\circledcirc$ and $\varepsilon_\circledcirc$ satisfy $(\varepsilon_\circledcirc)^\dagger = \sqrt{n}\,|a'_1\rangle$. It is straightforward to verify that this can be done such that $|b'_1\rangle = |b_1\rangle$, ensuring the coherence of $\mathscr{O}_\circ$ and $\mathscr{O}_\circledcirc$.

For this reason we will from now on assume that pairs of complementary observables are always coherent.

3.1 Strongly complementary observables

Many familiar observables, when expressed in terms of algebras, turn out to have useful additional properties. These are called *strongly complementary*; before describing them we will require some preliminary definitions.

Definition 13. A *(commutative) bialgebra* on X is a 4-tuple $(\mu, \eta, \delta, \varepsilon)$ of maps,

$$
\begin{aligned}
\mu &: X \otimes X \to X & \delta &: X \to X \otimes X \\
\eta &: I \to X & \varepsilon &: X \to I,
\end{aligned}
$$

which we write graphically as ($\hat{\mathcal{A}}$, $\hat{\circ}$, $\hat{\mathcal{Y}}$, $\hat{\circ}$), such that:

- (X, μ, η) is a (commutative) monoid;
- (X, δ, ε) is (cocommutative) comonoid; and,
- the following additional equations are satisfied:

$$\tag{12}$$

$$\tag{13}$$

$$\tag{14}$$

Remark 12. Note that equations (13) and (14) are very similar to the equations required for the coherence of two observables, per Definition 12. The only difference there is that the scalar $\begin{smallmatrix}\circ\\\circ\end{smallmatrix}$ is not assumed to be trivial.

Definition 14. A *(commutative) Hopf algebra* on X is a (commutative) bialgebra on X, augmented with a map $s : X \to X$, called the *antipode*, which satisfies:

$$\tag{15}$$

Again, note the similarity to Equation (11): the difference is only by a scalar factor.

Definition 15. A pair $(\mathscr{O}_\circ, \mathscr{O}_\circleddot)$ of observables on the same object X is *strongly complementary* iff they are coherent and:

$$\text{(diagram)} \tag{16}$$

To expand on this definition slightly, given a pair of strongly complementary observables, if we consider just the monoid part of one and the comonoid part of the other then the resulting structure is, up to a scalar factor, a bialgebra. Note that thanks to the up-down symmetry induced by the † it doesn't matter which is the monoid and which the comonoid. For obvious reasons, the we say that a pair of strongly complementary observables forms a *scaled bialgebra*, and we refer to Equation (16) as the *bialgebra law*. Notice that we have not, as yet, established any connection between complementarity (Definition 11) and strong complementarity. The following theorem links the two definitions.

Theorem 4. *Let $\mathscr{O}_\circ$ and $\mathscr{O}_\circleddot$ be strongly complementarity observables; then they are complementary.*

Proof. Let s be defined by

$$\text{(diagram)}$$

as per Equation (9). *Using the bialgebra law we reason:*

$$\text{(diagram)}$$

The last equation relies on the fact that $\eta_\circ$ is classical for $\mathscr{O}_\circleddot$ (and $\eta_\circleddot$ for $\mathscr{O}_\circ$), and (7).

As a consequence, strongly complementary observables always form a *scaled Hopf algebra*. Note that Theorem 4 relies on the fact that $\mathscr{O}_\circ$ and $\mathscr{O}_\circleddot$ are Frobenius algebras; it is certainly not the case that every bialgebra is a Hopf algebra.

The converse to Theorem 4 does not hold: it is possible to find coherent complementary observables in **FHilb** which are not strongly complementary. See [29] for a counterexample.

The following lemma about the antipode for a strongly complementary pair was shown in [32].

Lemma 1. *If $(\mathscr{O}_\circ, \mathscr{O}_\circleddot)$ are strongly complementary, and have enough classical points then the antipode s is:*

- *self-adjoint;*
- *a monoid homomorphism from $\mathscr{O}_\circ$ to itself, and similarly for $\mathscr{O}_\circleddot$; and*
- *a comonoid homomorphism from $\mathscr{O}_\circ$ to itself, and similarly for $\mathscr{O}_\circleddot$.*

3.2 Strong Complementarity and Phase Groups

For complementary observables, classical points of one observable are always included in the phase group of the other observable, up to a normalizing scalar. Strong complementarity strengthens this property to inclusion as a subgroup. Let $\mathscr{K}_\odot$ be the set of classical points of $\mathscr{O}_\odot$ multiplied by the scalar factor $\begin{smallmatrix}\bigcirc\\\bigcirc\end{smallmatrix}$.

Theorem 5. Let $(\mathscr{O}_\circ, \mathscr{O}_\odot)$ be strongly complementary observables and let $\mathscr{O}_\odot$ have finitely many classical points. Then $\mathscr{K}_\odot$ forms a subgroup of the phase group $\Phi_\circ$ of $\mathscr{O}_\circ$. The converse also holds when $\mathscr{O}_\circ$ has 'enough classical points'.

Proof. By strong complementarity it straightforwardly follows that, up to a scalar, $\mu_\circ$ applied to two classical points of $\mathscr{O}_\odot$ yields again a classical point of $\mathscr{O}_\odot$:

The unit of $\Phi_\circ$ is, up to a scalar, also a classical point of $\mathscr{O}_\odot$ by coherence. Hence, $\mathscr{K}_\odot$ is a submonoid of $\Phi_\circ$ and any finite submonoid is a subgroup. The converse follows from:

together with the 'enough classical points' assumption.

Recall that the exponent of a group G is the maximum order of any element of that group: $\exp(G) = \max\{|g| : g \in G\}$.

Corollary 1. For any pair of strongly complementary observables, let $k = \exp(\mathscr{K}_\odot)$. Then, assuming $\mathscr{O}_\circ$ has 'enough classical points':

$$(17)$$

Proof. In a finite abelian group, the order of any element divides $\exp(\mathscr{K}_\odot)$. The result then follows by:

together with the 'enough classical points' assumption.

Proposition 4. For a pair of strongly complementary observables $\overset{i}{\underset{}{\bigcirc}}$ is a $\mathscr{O}_\circ$-homomorphism for all $\overset{i}{\underset{}{\bigcirc}} \in \mathscr{K}_\circ$. Conversely, this property defines strong complementarity provided $\delta_\circ$ has 'enough classical points'.

Proof. Similar to the proof of Thm. 5.

3.3 Classification of Strong Complementarity in **FHilb**

Corollary 2. Every pair of strongly complementary observables in **FHilb** is of the following form:

$$\begin{cases} \delta_\circ :: |g\rangle \mapsto |g\rangle \otimes |g\rangle \\ \varepsilon_\circ :: |g\rangle \mapsto 1 \end{cases} \quad \begin{cases} \delta_\circ^\dagger :: |g\rangle \otimes |h\rangle \mapsto \frac{1}{\sqrt{D}}|g+h\rangle \\ \varepsilon_\circ^\dagger :: 1 \mapsto \sqrt{D}|0\rangle \end{cases}$$

where $(G = \{g, h, \ldots\}, +, 0)$ is a finite Abelian group. Conversely, each such pair is always strongly complementary.

Proof. By Theorem 5 it follows that the classical points of one observable (here $\mathscr{O}_\circ$) form a group for the multiplication of the other observable (here $\delta_\circ^\dagger$), and in **FHilb** this characterises strong complementarity.

One of the longest-standing open problems in quantum information is the characterisation of the number of pairwise complementary observables in a Hilbert space of dimension D. In all known cases this is $D+1$, and the smallest unknown case is $D = 6$. We now show that in the case of strong complementarity this number is always 2 for $D \geq 2$.

Theorem 6. In a Hilbert space with $D \geq 2$ the largest set of pairwise strongly complementary observables has size at most 2.

Proof. Assume that both $(\mathscr{O}_\circ, \mathscr{O}_\circ)$ and $(\mathscr{O}_\circ, \mathscr{O}_\bullet)$ are strongly complementary pairs. By coherence $\overset{}{\underset{}{\bigcirc}}$ and $\overset{}{\underset{}{\bullet}}$ must be proportional to classical points of $\mathscr{O}_\circ$. If $(\mathscr{O}_\circ, \mathscr{O}_\bullet)$ were to be strongly complementary observables, it is easily shown that $\overset{\bullet}{\underset{\bigcirc}{}} \neq 0$ so $\overset{}{\underset{}{\bigcirc}}$ and $\overset{}{\underset{}{\bullet}}$ are proportional to the same classical point. Hence, up to a non-zero scalar:

$$\Big\vert \; = \; \overset{\bullet}{\searrow_\bullet} \; = \; \overset{\bullet}{\searrow_\circ} \; = \; \overset{\bigcirc}{\underset{\bigcirc}{}}$$

i.e. the identity has rank 1, which fails for $D \geq 2$. By Corollary 2 a strongly complementary pair exists for any $D \geq 2$.

4 Mixed states, measurements, and "abstract probabilities"

For some ket $|\psi\rangle$ in a Hilbert space, there are (at least) four distinct ways to represent $|\psi\rangle$ as a linear map.

It is possible to represent a ket $|\psi\rangle \in H$ as a map $|\psi\rangle : \mathbb{C} \to H$, sending $1 \in \mathbb{C}$ to $|\psi\rangle$. We call such a map a "point" of H, because it does nothing more than picking out a specific element. The second map is the associated bra $\langle\psi| : H \to \mathbb{C}$. This kind of map is called a "co-point". We can also regard such a map as an element of the dual space H^*. But then, H^* is just another Hilbert space, so we could just as well represent $\langle\psi|$ as a point of H^*. That is, a linear map $\langle\psi|^* : \mathbb{C} \to H^*$. There is yet a fourth way to represent $|\psi\rangle$, namely as a linear map $|\psi\rangle^* : H^* \to \mathbb{C}$, sending a bra $\langle\phi| \in H^*$ to the inner product $\langle\phi|\psi\rangle \in \mathbb{C}$.

So, for a given ket $|\psi\rangle$, there are four ways to write it as points or copoints.

The difference in these four pictures is largely notational: the data they represent is the same. However, its a very useful piece of notation, especially when representing functionals between spaces of maps. Firstly, we note that we can represent a map $M : A \to B$ as a special kind of point, $|\Psi_M\rangle \in A^* \otimes B$.

$$M = \sum a_i^j \, |j\rangle \langle i| \qquad \mapsto \qquad |\Phi_M\rangle = \sum a_i^j \, \langle i| \otimes |j\rangle$$

These two objects clearly represent the same data. In fact, this mapping is essentially the Choi-Jamiołkowski isomorphism. However instead of fixing a basis, we rely on the dual space A^*. Thus the value on the right does not depend on a choice of orthonormal basis. By fixing a basis $\mathscr{B} = \{|i\rangle\}$, we can define a transposition map $T_{\mathscr{B}}(|i\rangle) = \langle i|$. Then the usual C-J isomorphism is recovered as $(T_{\mathscr{B}} \otimes 1)|\Phi_M\rangle$. However, this *does* depend on a choice of basis, since $T_{\mathscr{B}}$ does.

In [8], the authors refer to $|\Psi_M\rangle$ as the "name" of a map. We shall often find this representation more useful than the usual C-J representation, especially in instances involving several distinct orthonormal bases. Using the caps and cups from before, we can isomorphically relate maps and their associated names.

The benefit of working with names of maps, as opposed to the maps themselves becomes clear when we start considering higher-order functionals. For a finite-dimensional Hilbert space H, let $L(H)$ be the space of linear maps from H to itself. When operating on density matrices, we often want to consider maps of the form

$\Phi\colon L(H) \to L(K)$. We can either treat this as a genuine, higher-order map, or we can treat it as a first-order map from names to names.

$$\left(\boxed{\rho}\right) \quad \Rightarrow \quad \begin{array}{c} \boxed{\Phi} \\ \bigtriangledown\!\!\!\rho \end{array}$$
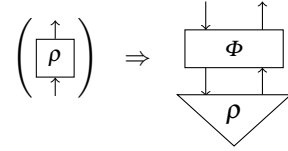
Since, in finite dimensions, we have an isomorphism $L(H) \cong H^* \otimes H$, we know that all maps $\Phi\colon L(H) \to L(K)$ can be represented this way.

In ordinary quantum theory, mixed quantum states are represented as positive operators and operations as completely positive maps, or CPMs. These are maps that take positive operators to positive operators. A general CPM can be written in terms of a set of linear maps $\{B_i : H \to K\}$ called its *Kraus maps*.

$$\Theta(\rho) = \sum_i B_i \rho B_i^\dagger$$

As before, we can collapse the higher-order map Θ to a first- order map by translating the positive operator ρ to its associated name.

$$\bigtriangledown\!\!\!\rho \quad := \quad \boxed{\rho}$$
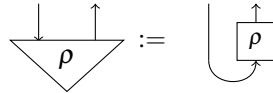
Then, we can encode the Kraus vectors of Θ in a map $B' = \sum |i\rangle \otimes B_i$ and represent Θ as:

$$\boxed{B'_*} \quad \boxed{B'} \tag{18}$$
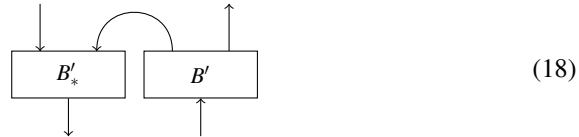
When we take the elements in Eq. (19) to be morphisms in an arbitrary †-compact category, this gives us an abstract definition of a completely positive map. This is Selinger's representation of CPMs [33].

$$\boxed{f_*} \quad \boxed{f} \tag{19}$$
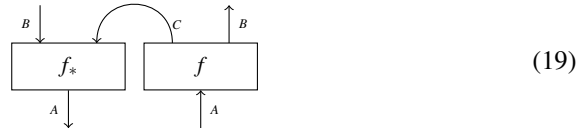
Important special cases are *states* where $A \cong I$, *effects* where $B \cong I$, and 'pure' maps, where $C \cong I$.
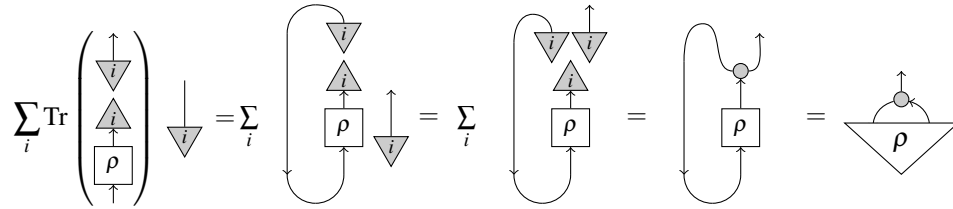
4.1 Measurements and Born vectors

Returning to quantum mechanics, we can see how a quantum measurement would look in this language. A (projective) quantum measurement $M_\circ$ is a CPM that sends trace 1 positive operators (in this case quantum states) to trace 1 positive operators that are diagonal in some ONB (encoding a probability distribution of outcomes). Suppose we wish to measure with respect to an observable $\mathscr{O}_\circ$, whose classical points form an ONB $\{|x_i\rangle\}$. The probability of getting the *i*-th measurement outcome is computed using the Born rule.

$$\text{Prob}(i, \rho) = \text{Tr}(|x_i\rangle\langle x_i|\rho)$$

We can write this probability distribution as a vector in the basis $\{|x_i\rangle\}$. That is, a vector whose *i*-th entry is the probability of the *i*-th outcome:

$$M_\circ(\rho) = \sum \text{Tr}(|x_i\rangle\langle x_i|\rho)\,|x_i\rangle$$

So, *M* defines a linear map from density matrices to probability distributions. Expanding this graphically, we have:
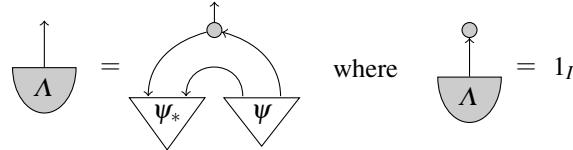

We are now ready to make definitions for abstract measurements and abstract probability distributions, which we shall call Born vector.

**Definition 16.** For an observable structure $\mathscr{O}_\circ$, a measurement is defined as the following map:

Any point $|\Lambda)_\circ : I \to X$ of the following form is called a *Born vector*, with respect to $\mathscr{O}_\circ$:

Theorem 7. *In* **FHilb***, Born vectors for an observable $\mathscr{O}_\circ$ are precisely those vectors whose entries are positive and sum to* 1*, when written in the basis of $\mathscr{O}_\circ$.*

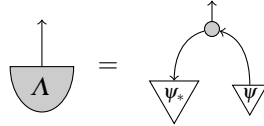*Proof.* Let $|\Lambda)_\circ$ be a Born vector. Its *i*-th coefficient in the $\mathscr{O}_\circ$-basis is given by:
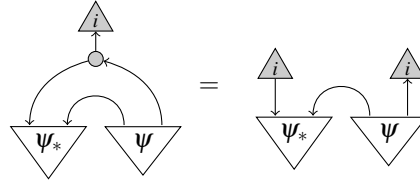
We can see that these coefficients sum to 1 by using the definition of the deleting point:

This is a completely positive map from $\mathbb{C}$ to $\mathbb{C}$. In other words, it is a positive scalar. For the converse, assume $|\Lambda)_{\circ}$ is a probability distribution whose i-th coefficient in

the $\mathscr{O}_{\circ}$-basis is $p_i \in \mathbb{R}_+$. Then, letting $\psi = \sum \sqrt{p_i}$:

Post-composing with the deleting point yields $\sum(\sqrt{p_i})^2 = \sum p_i = 1$.

Thus Born vectors in **FHilb** correspond precisely to *probability distributions* over classical points.

We can naturally extend the definition above to points of the form $|\Lambda)_{\circ} : I \to X \otimes \ldots \otimes X$ by requiring that they be Born vectors with respect to the product Frobenius algebra $\mathscr{O}_{\circ} \otimes \ldots \otimes \mathscr{O}_{\circ}$.

The adjoint of the measurement map $m_{\circ}^{\dagger}$ is a preparation operation. In **FHilb**, it takes a Born vector $|\Lambda)_{\circ}$ with respect to $\mathscr{O}_{\circ}$ and produces a probabilistic mixture of the (pure) outcome states of $\mathscr{O}_{\circ}$ with probabilities given by $|\Lambda)_{\circ}$.

This leads to a simple classical vs. quantum diagrammatic paradigm that applies in all of the models we consider [22]: *classical systems are encoded as a single wire and quantum systems as a double wire*. The same applies to operations, and $m_{\circ}$ and $m_{\circ}^{\dagger}$ allow passage between these types.

Note that the classical data will 'remember' to which observable it relates, cf. the encoding $\sum_i p_i |x_i\rangle$. This is physically meaningful since, for example, when one measures position the resulting value will carry specification of the length unit in which it is expressed. If one wishes to avoid interconversion of this 'classical data with memory', one could fix one observable, and unitarily transform the quantum data before measuring. Indeed, if
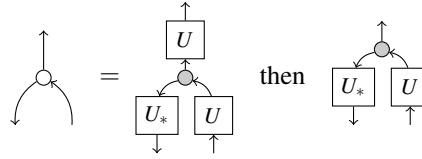
measures the $\mathcal{O}_\circ$-observable but produces $\mathcal{O}_\bullet$-data. In **FHilb**, all observable structures are unitarily isomorphic, so any projective measurement can be obtained in this way. A particularly relevant example is when these unitaries are phases with respect the another observable structure $\mathcal{O}_\circ$.

$$m_\bullet^\alpha := \quad \text{} \tag{20}$$

When $\mathcal{O}_\circ$ is induced by the Pauli spin-Z observable and $\mathcal{O}_\bullet$ by the Pauli spin-X observable, then $m_\bullet = m_\bullet^0$ is an X measurement and $m_\bullet^{\pi/2}$ is a Y measurement. Note however, that both produce Born vectors of outcome probabilities with respect to the $\mathcal{O}_\bullet$ basis. This will be useful in the next section.

5 Example: non-locality of QM

In 1989 Greenburger, Horne, and Zeilinger provided an analysis [34] of quantum theory which improves on Bell's theorem in one crucial way. Whereas Bell demonstrated a *probabilistic* argument that quantum theory is incompatible with the assumption of local realism (i.e. quantum theory generates correlations that are too high for a classical local hidden variable model), the GHZ/Mermin theorem illustrates a situation that rules out a locally realistic model *possibilistically*. That is, they described a series of experiments for which quantum theory predicts a single, definite outcome that is impossible under the assumption of locality.

Here, we reproduce Mermin's version of this argument [35] using diagrammatic techniques. There are two key ingredients for this translation. The first is a graphical notion of locality. For our purposes, it will suffice to treat locality as the fact that global probability can be traced down to hidden states that determine all measurement outcomes, since we shall show that no hidden state can ever be compatible with the predictions of quantum theory. Hence, there is no point in even considering crafting a local hidden variable representation.

The second key ingredient is *parity*. The GHZ/Mermin trick for producing definite outcomes is to look not at individual measurement outcomes, but the overall parity of those outcomes, i.e. "Did the experiment produce and even or an odd number of clicks?". Considering a single outcome (click or no-click) as an element of the abelian group $\mathbb{Z}_2$, the parity of a set of outcomes is given by their sum in the group. We already saw in section 3.3 that strongly complementary observables are classified by abelian groups. In two dimensions, there is only one such strongly

complementary pair, namely the one corresponding to $\mathbb{Z}_2$. When we prepare a GHZ state with respect to a certain observable (e.g. spin-Z) and conduct measurements using a strongly complementary observable (e.g. spin-X), we will see this $\mathbb{Z}_2$ structure arise.

By combining these two elements (the topological picture of locality and the encoding of abelian groups as strongly complementary observables) we will derive a contradiction. Furthermore in section 5.4, we shall see how strong complementary was embedded in the pre-conditions of a GHZ/Mermin-style argument in the first place.

5.1 A local hidden variable model

For a particular *n*-party state $|\Psi\rangle$ in some theory, a *local hidden variable (LHV)* model for that state consists of:

- a family of hidden states $|\lambda\rangle$, each of which assigns to any measurement on each subsystem a definite outcome; and,
- a probability distribution on these hidden states,

which simulates the probabilities of that theory. We say that a theory is *local* if each state admits a LHV model.

Consider three systems and four possible (compound) measurement settings, XXX, XYY, YXY, and YYX. A hidden state of an underlying LHV model stores one measurement outcome for each setting on each system:

$$|\lambda'\rangle = | \underbrace{\overbrace{+}^{X} \overbrace{-}^{Y}}_{system\ 1} \underbrace{\overbrace{-}^{X} \overbrace{+}^{Y}}_{system\ 2} \underbrace{\overbrace{-}^{X} \overbrace{+}^{Y}}_{system\ 3} \rangle$$

We can represent this diagrammatically as follows:
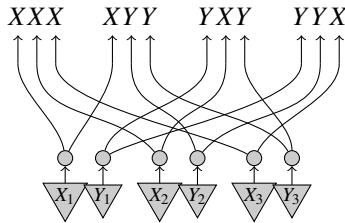
that is, we simply copy those values to each of the four measurement settings.

## 5.2 Encoding the GHZ state and computing correlations, diagrammatically

To present Mermin/GHZ style argument graphically, we first show how to compute measurement outcomes for an *n*-party GHZ state graphically. This computation relies on a standard theorem about bialgebras, which relates a graph-theoretic property of diagrams to equality of bialgebra expressions.

Definition 17. Let $(\,\raisebox{-2pt}{\scriptsize⚇},\,\raisebox{-2pt}{\scriptsize⊙},\,\raisebox{-2pt}{\scriptsize⚥},\,\raisebox{-2pt}{\scriptsize♀}\,)$ be a commutative, cocommutative bialgebra, and let $\mathscr{D}$ be a diagram consisting only of $\,\raisebox{-2pt}{\scriptsize⚇},\,\raisebox{-2pt}{\scriptsize⊙},\,\raisebox{-2pt}{\scriptsize⚥},\,♀$, identity maps, and swaps. Then, the characteristic matrix χ of $\mathscr{D}$ is a matrix of natural numbers where the (i,j)-th entry represents the number of forward-directed paths connecting the i-th input to the j-th output.

*Example 21.* The following terms have characteristic matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} :$$

The following terms have characteristic matrix
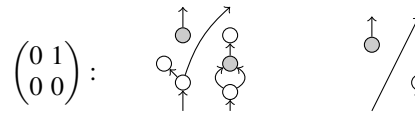
$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} :$$
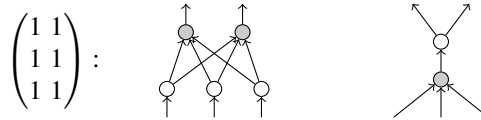
Theorem 8. *If two diagrams generated by the same bialgebra have the same characteristic matrix, they are equal as maps.*

Proof. (sketch) It is possible to show by case analysis that the three bialgebra equations can be used to move all of the gray dots to the top all the white dots to the bottom.
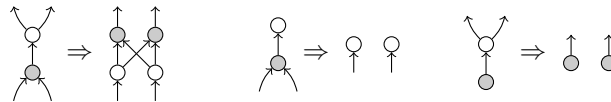
We can furthermore show that all three of these transformations preserve the characteristic matrix of $\mathscr{D}$. Once this is done, we obtain a diagram in normal form:

$$\leftrightarrow \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$
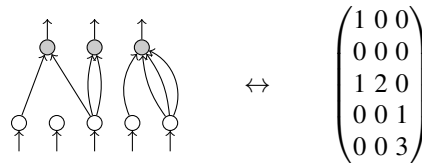
Then, it is possible to show there is *exactly one* such normal form for each characteristic matrix. In fact, it is straightforward to read off the matrix by counting edges in the normal form. Since every diagram can be put into normal form using equations that preserve the characteristic matrix, and normal forms are in 1-to-1 correspondence with characteristic matrices, this completes the proof.[9]

We can now apply the theorem to prove the following corollary.

Corollary 3. *The following equation holds for any connected bipartite graph with directions as shown.*

$$\text{(21)}$$

Proof. For the theorem on bialgebras to apply, all of the edges need to be directed upward. For a strongly complementary observable, the edge direction between two different colours can be reversed by applying the antipode S. Then, we use the fact that S is a Frobenius algebra endomorphism to move it down.

We apply Theorem 8 and the spider theorem to complete the proof.

We compute the classical probability distributions (= $\mathcal{O}_\circ$-data) for n measurements against arbitrary phases $\alpha_i \in \Phi_\circ$ on n systems of any type in a generalised $GHZ_\circ^n$-state:

$GHZ_\circ^n := (\circ)_0^n$

---

[9] For a formal statement and proof of this theorem, in terms of factorisation systems see [36].

Applying Corollary 3, we note that this is a probability distribution followed by a $\bigcirc$-copy.

$$(*) = \qquad\qquad =: \qquad\qquad (22)$$

The following is an immediate consequence.

Theorem 9. When measuring each system of a GHZ_A^n-state against an arbitrary angle then the resulting classical probability distribution over outcomes is symmetric.

Theorem 10. The classical probability distributions for $m_{\circ}^{\alpha_1} \otimes \ldots \otimes m_{\circ}^{\alpha_n}$-measurements on a GHZ_A^n-state is:

- uncorrelated if $|\sum \alpha_i\rangle_{\circ}$ is a classical point for $\mathscr{O}_{\circ}$ and,
- parity-correlated if $|\sum \alpha_i\rangle_{\circ}$ is a classical point i for $\mathscr{O}_{\circ}$ (i.e. contains precisely those outcomes $i_1 \otimes \ldots \otimes i_n$ such that the sum of group elements $\sum i_k$ is equal to i).

Example 22. We can compute the outcome distributions for XXX, XYY, YXY, and YYX measurements on three qubits in a GHZ-state using the technique described above. First, outcome distribution $|A\rangle_{\circ}$ for XXX:
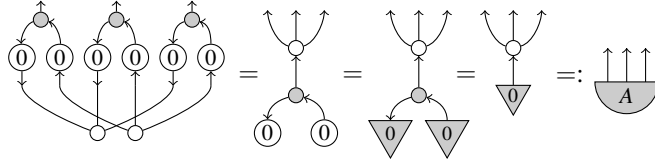
Next, we compute outcome distribution $|B_1\rangle_{\circ}$ for XYY:
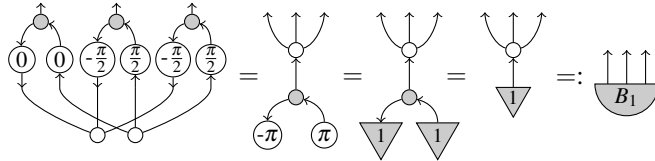
Computing correlations as in Figure (22) is symmetric in the choice of measurement angle for each of the systems. Thus, for the other two cases (YXY and YYX), we get the same result: $|B_1\rangle_{\circ} = |B_2\rangle_{\circ} = |B_3\rangle_{\circ}$.

5.3 Deriving the contradiction

Consider the function:

(23)

We have already seen that strongly complementary observables correspond to group algebras. That is, defines a group algebra over the classical points of $\mathscr{O}_\odot$. For qubits there is only one choice: $\mathbb{Z}_2$. Thus, this function computes the parity (i.e. the $\mathbb{Z}_2$-sum) of all outcomes.

Measuring the parity for any local hidden state we obtain:
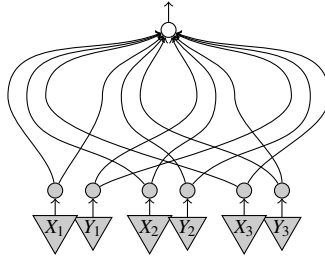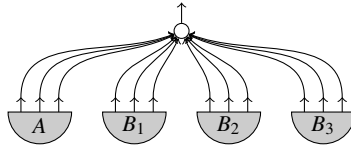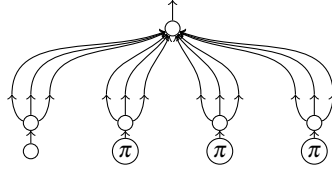

that is, by (17):
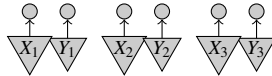
and hence:

and measuring the parity in quantum theory we obtain:

that is, by the previous section:

and hence:

which yields a contradiction.

## 5.4 GHZ/Mermin assumptions and the necessity of strong complementarity

We shall examine two assumptions that play a key role in a GHZ/Mermin style non-locality argument, and show that the presence of a strongly complementary observable arises as a consequence.

The original argument due to Greenburger, Horne, and Zeilinger [34] and later simplifications [37, 35] focus on a state defined in terms of correlated (or anti-correlated) Z-spins and local spin measurements in the XY-plane. We generalise this assumption as follows.

Assumption 1 (Coherence).   We will use a GHZ state defined with respect to some observable structure $\mathscr{O}_\circ$. Measurements are all conducted within a $\mathscr{O}_\circ$-phase of some coherent observable $\mathscr{O}_\circledcirc$.

In **FHilb**, all observables containing at least one unbiased classical point, w.r.t. $\mathscr{O}_\circ$, are within a $\mathscr{O}_\circ$-phase of a coherent observable, so we could weaken this assumption further. That is, if $\mathscr{O}_\circledcirc$ contains an unbiased classical point, we might as well assume it is coherent, since **Assumption 1** allows us to freely choose phases.

By **Assumption 1**, the correlations associated with each experiment are computed from this diagram:
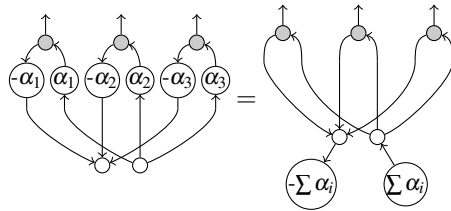
The second assumption is what [34] refers to as "super-classicality". We shall refer to it as *sharpness*.

Assumption 2 (Sharpness).   After all subsystems except one are measured, the final measurement outcome is fixed.

The map is called the *decoherence map* for $\mathcal{O}_\circ$. It projects from the space of all quantum mixed states to the the space of classical mixtures of eigenstates of $\mathcal{O}_\circ$. To assert sharpness, we require that, once two of the three systems are measured, the third is invariant under this map:
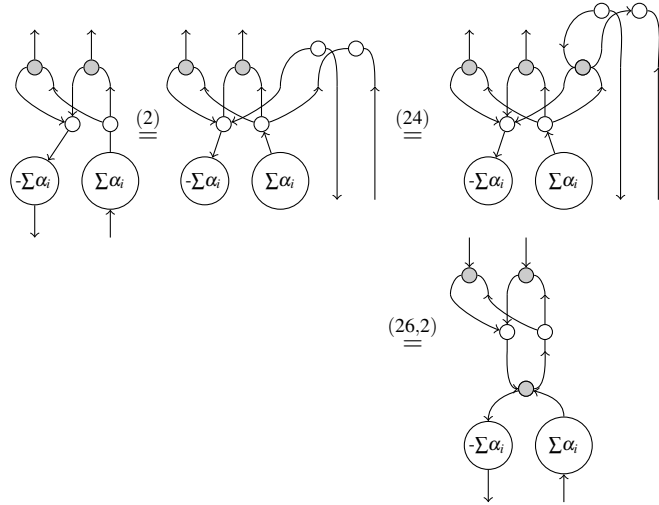

$$\tag{24}$$

Plugging the unit of $\mathcal{O}_\circ$ in the 2nd system both for LHS and RHS, and using coherence we obtain:

$$\tag{25}$$

and by exploiting symmetry we have:

$$\tag{26}$$

Hence we obtain:

Since $\delta_\circ^\dagger \circ (1_X \otimes \sum_i \alpha_i)$ is unitary it cancels. Thus our assumptions lead us to conclude the following equation for the observable structures $(\mathscr{O}_\circ, \mathscr{O}_\bullet)$:
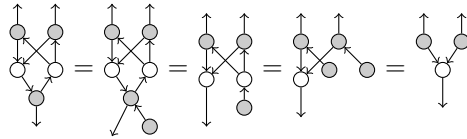
$$\tag{27}$$

Proposition 5. A pair $(\mathscr{O}_\circ, \mathscr{O}_\bullet)$ of coherent observables satisfying Equation (27) are strongly complementary.
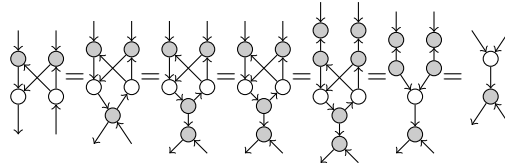
Proof. First, we show that equation (27) implies the following, for any pair of coherent observables:
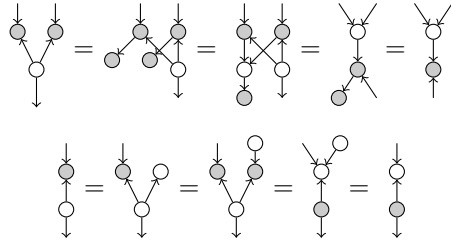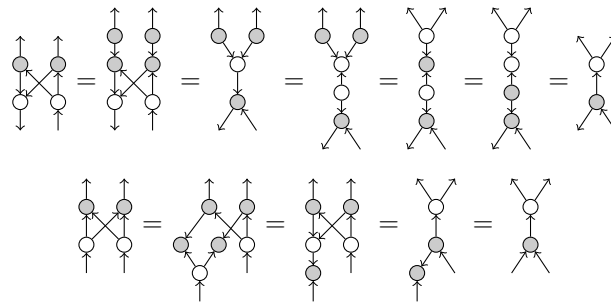
$$\tag{28}$$

The proof goes as follows:

...which implies:

Equation (28) is very nearly the required equation for strong complementarity, but the directions are wrong. However, we can correct this by first showing the following equations, using coherence and (28):

Then, we complete the proof by using the equations above to change the directions of the arrows on the inside:

Thus any coherent pair of observable structures satisfying Equation (27) is a strongly complementary pair.

6 Summary and Further Reading

In this chapter, we developed the notion of a *generalised compositional theory*, a new approach to studying quantum mechanics and constructing foil theories with quantum-like properties. The main building blocks for a GCT are:

- a collection of systems $A, B, C, \ldots$,
- a collection of *primitive* processes, and
- a notion of horizontal composition $\otimes$ and vertical composition $\circ$.

From this sparse setting, we began to add extra pieces of structure.

- symmetry maps $\Rightarrow$ "permutibility of systems"
- dagger $\Rightarrow$ "time-reversed processes"
- duals $\Rightarrow$ "map/state duality"

This structure and its diagrammatic presentation give a rich language for talking about composed processes. We then went on to define various concepts in this

framework, often by analogy to their Hilbert space counterparts: pure states, reversible dynamics, quantum observables, complementarity, mixed states, and measurements. Using these ingredients, we worked through a complete example, following Mermin's illustration of a possibilistic locality violation, as predicted by quantum mechanics.

The interested reader can find many papers related to, or extending the formalism introduced in this chapter. One example is the *ZX-calculus*, which is a graphical calculus for the interaction of the Pauli-Z and Pauli-X observable structures. In addition to the usual rules (complementarity, strong complementarity), several other rules are added, which turn out to be complete for stabiliser quantum mechanics [38]. This calculus has been applied to the study of measurement-based quantum computing [15], topological MBQC [16], and quantum protocols [39].

The ideas developed in section 4 originated in [22]. In [40], a simplified formalism for interacting classical and quantum data was developed, and can be viewed as an abstraction of the C*-algebraic approach to the study of quantum information.

References

1. J. Barrett, Physical Review A **75**(032304) (2007)
2. M. Pawlowski, T. Paterek, D. Kazlikowski, V. Scarani, A. Winter, M. Zukowski, Nature **461**(1101) (2009). arXiv:0905.2292
3. H. Barnum, J. Barrett, L.O. Clark, M. Leifer, R.W. Spekkens, N. Stepanik, A. Wilce, R. Wilke, New Journal of Physics **12**(033024) (2009). arXiv:0909.5075
4. E. Schrödinger, in *Proceedings of the Cambridge Philosophical Society*, vol. 31 (Academic Press, 1935), pp. 555–563
5. R. Penrose, in *Combinatorial Mathematics and its Applications* (Academic Press, 1971), pp. 221–244
6. G.M. Kelly, M.L. Laplaza, Journal of Pure and Applied Algebra **19**, 193 (1980)
7. A. Joyal, R. Street, Advances in Mathematics **102**, 20 (1993)
8. S. Abramsky, B. Coecke, in *Proceedings of 19th IEEE conference on Logic in computer science* (IEEE Press, 2004), LiCS'04, pp. 415–425
9. B. Coecke, in *Quantum Theory: Reconsiderations of the Foundations III* (AIP Press, 2005), pp. 81–98
10. G.M. D'Ariano, G. Chiribella, P. Perinotti, Physical Review A **84**(012311) (2010). arXiv:1011.6451
11. L. Hardy, in *Deep Beauty: Understanding the Quantum World through Mathematical Innovation* (Cambridge University Press, 2011), pp. 409–442. arXiv:0912.4740
12. B. Coecke, B. Edwards, R.W. Spekkens, Electronic Notes in Theoretical Computer Science **270**(2), 15 (2011)
13. B. Edwards, Phase groups and local hidden variables. Tech. Rep. RR-10-15, Dept. of computer science, University of Oxford (2010)
14. B. Coecke, A. Kissinger, in *Lecture Notes in computer science*, vol. 6199 (Springer, 2010), pp. 297–308
15. R. Duncan, S. Perdrix, in *Proceedings of the 37th international colloquium conference on Automata, languages and programming: Part II* (Springer-Verlag, Berlin, Heidelberg, 2010), ICALP'10, pp. 285–296. URL http://dl.acm.org/citation.cfm?id=1880999.1881030
16. C. Horsman, New Journal of Physics **13**(095011) (2011). arXiv:1101.4722
17. S. Mac Lane, *Categories for the Working Mathematician (2nd Ed.)* (Springer-Verlag, 1997)

18. B. Coecke, E.O. Paquette, in *New Structures for Physics*, *Springer Lecture Notes in Physics*, vol. 813 (2011), pp. 173–286
19. P. Selinger, in *New Structures for Physics*, *Springer Lecture Notes in Physics*, vol. 813 (2011), pp. 289–355
20. Y. Lafont, Journal of Pure and Applied Algebra **184**(2-3), 257 (2003)
21. A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J.A. Smolin, H. Weinfurter, Phys. Rev. A **52**, 3457 (1995). DOI 10.1103/PhysRevA.52.3457
22. B. Coecke, E.O. Paquette, D. Pavlovic, in *Semantic Techniques for Quantum Computation* (Cambridge University Press, 2009), pp. 29–69
23. D.G.B.J. Dieks, Physics Letters A **92**, 271 (1982)
24. W.K. Wootters, W. Zurek, Nature **299**, 802 (1982)
25. A.K. Pati, S.L. Braunstein, Nature **404**, 164 (2000). arXiv:quant-ph/9911090
26. B. Coecke, D. Pavlovic, J. Vicary, Mathematical Structures in Computer Science **23**, 555 (2013)
27. D. Pavlovic, in *Lecture Notes in computer science*, vol. 5494 (Springer, 2009), pp. 143–157. arXiv:0812.2266
28. B. Coecke, R. Duncan, in *Lecture Notes in computer science*, vol. 5126 (Springer, 2008), pp. 298–310
29. B. Coecke, R. Duncan, New Journal of Physics **13**(043016) (2011). arXiv:0906.4725
30. R.W. Spekkens, Physical Review A **75**(032110) (2007). arXiv:quant-ph/0401052
31. B. Coecke, B. Edwards. Spekkens's toy theory as a category of processes. arXiv:1108.1978v1 [quant-ph] (2011)
32. A. Kissinger, Pictures of processes: Automated graph rewriting for monoidal categories and applications to quantum computing. Ph.D. thesis, University of Oxford (2012)
33. P. Selinger, Electronic Notes in Theoretical Computer Science **170**, 139 (2007)
34. D.M. Greenberger, M.A. Horne, A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, ed. by M. Kafatos (Springer, 1989), pp. 69–72
35. N.D. Mermin, American Journal of Physics **58**, 731 (1990)
36. S. Lack, Theory and Applications of Categories **13**(9), 147 (2004)
37. M.A. Horne, A. Shimony, D M Greenberger, A. Zeilinger, American Journal of Physics **58**, 1131 (1990)
38. M. Backens, in *Proceedings of Quantum Physics and Logic* (2012), pp. 15–27
39. A. Hillebrand, Quantum protocols involving multiparticle entanglement and their representations in the zx-calculus. Master's thesis, University of Oxford (2011)
40. B. Coecke, C. Heunen, A. Kissinger, in *Proceedings of Quantum Physics and Logic* (2012), pp. 87–100