Categories and Quantum Informatics: Monoidal categories

Chris Heunen

Spring 2018

A monoidal category is a category equipped with extra data, describing how objects and morphisms can be combined 'in parallel'. This chapter introduces the theory of monoidal categories, and shows how our example categories **Hilb**, **Set** and **Rel** can be given a monoidal structure. We also introduce a visual notation called the graphical calculus, which provides an intuitive and powerful way to work with them.

1.1 Monoidal structure

Throughout this book, we interpret objects of categories as systems, and morphisms as processes. A monoidal category has additional structure allowing us to consider processes occurring *in parallel*, as well as sequentially. In terms of our example categories from the introduction, one could interpret this in the following ways:

- letting independent physical systems evolve simultaneously;
- running computer algorithms in parallel;
- taking products or sums of algebraic or geometric structures;
- using separate proofs of P and Q to construct a proof of the conjunction (P and Q).

It is perhaps surprising that a nontrivial theory can be developed at all from such simple intuition. But in fact, some interesting general issues quickly arise. For example, let A, B and C be processes, and write $\otimes$ for the parallel composition. Then what relationship should there be between the processes $(A \otimes B) \otimes C$ and $A \otimes (B \otimes C)$? You might say they should be equal, as they are different ways of expressing the same arrangement of systems. But for many applications this is simply too strong: for example, if A, B and C are Hilbert spaces and $\otimes$ is the usual tensor product of Hilbert spaces, these two composite Hilbert spaces are *not* exactly equal; they are only isomorphic. But we then have a new problem: what equations should that isomorphism satisfy? The theory of monoidal categories is formulated to deal with these issues.

Definition 1.1. A *monoidal category* is a category $\mathbf{C}$ equipped with the following data:

- a *tensor product* functor $\otimes \colon \mathbf{C} \times \mathbf{C} \to \mathbf{C}$;
- a *unit object* $I \in \mathrm{Ob}(\mathbf{C})$;
- an *associator* natural isomorphism $(A \otimes B) \otimes C \xrightarrow{\alpha_{A,B,C}} A \otimes (B \otimes C)$;
- a *left unitor* natural isomorphism $I \otimes A \xrightarrow{\lambda_A} A$;
- a *right unitor* natural isomorphism $A \otimes I \xrightarrow{\rho_A} A$.

This data must satisfy the *triangle* and *pentagon* equations, for all objects A, B, C and D:

$$\begin{tikzcd}
(A \otimes I) \otimes B \arrow[rr, "\alpha_{A,I,B}"] \arrow[dr, "\rho_A \otimes \mathrm{id}_B"'] & & A \otimes (I \otimes B) \arrow[dl, "\mathrm{id}_A \otimes \lambda_B"] \\
& A \otimes B &
\end{tikzcd} \tag{1.1}$$

$$\begin{tikzcd}
& \big(A \otimes (B \otimes C)\big) \otimes D \arrow[r, "\alpha_{A,B\otimes C,D}"] & A \otimes \big((B \otimes C) \otimes D\big) \arrow[dr, "\mathrm{id}_A \otimes \alpha_{B,C,D}"] & \\
\big((A \otimes B) \otimes C\big) \otimes D \arrow[ur, "\alpha_{A,B,C} \otimes \mathrm{id}_D"] \arrow[drr, "\alpha_{A\otimes B,C,D}"'] & & & A \otimes \big(B \otimes (C \otimes D)\big) \\
& & (A \otimes B) \otimes (C \otimes D) \arrow[ur, "\alpha_{A,B,C\otimes D}"'] &
\end{tikzcd} \tag{1.2}$$

The naturality conditions for α, λ and ρ correspond to the following equations:

$$\begin{tikzcd}
(A \otimes B) \otimes C \arrow[r, "\alpha_{A,B,C}"] \arrow[d, "(f \otimes g) \otimes h"] & A \otimes (B \otimes C) \arrow[d, "f \otimes (g \otimes h)"'] \\
(A' \otimes B') \otimes C' \arrow[r, "\alpha_{A',B',C'}"] & A' \otimes (B' \otimes C')
\end{tikzcd}
\qquad
\begin{tikzcd}
I \otimes A \arrow[r, "\lambda_A"] \arrow[d, "I \otimes f"] & A \arrow[d, "f"] & A \otimes I \arrow[l, "\rho_A"'] \arrow[d, "f \otimes I"'] \\
I \otimes B \arrow[r, "\lambda_B"'] & B & B \otimes I \arrow[l, "\rho_B"]
\end{tikzcd} \tag{1.3}$$

The tensor unit object I represents the 'trivial' or 'empty' system. This interpretation comes from the unitor isomorphisms λ_A and ρ_A, which witness the fact that the object A is 'just as good as', or isomorphic to, the objects $A \otimes I$ and $I \otimes A$.

The triangle and pentagon equations each say that two particular ways of 'reorganizing' a system are equal. Surprisingly, this implies that *any* two 'reorganizations' are equal; this is the content of the Coherence Theorem.

Theorem 1.2 (Coherence for monoidal categories)**.** *Given the data of a monoidal category, if the pentagon and triangle equations hold, then any well-typed equation built from α, λ, ρ and their inverses holds.* □

In particular, the triangle and pentagon equation together imply $\rho_I = \lambda_I$. To appreciate the power of the coherence theorem, try to show this yourself.

Coherence is the fundamental motivating idea of a monoidal category, and gives an answer to question we posed earlier in the chapter: the isomorphisms should satisfy *all* possible well-typed equations. So while these morphisms are not trivial—for example, they are not necessarily identity morphisms—it doesn't matter how we apply them in any particular case.

Our first example of a monoidal structure is on the category **Hilb**.

Definition 1.3. The monoidal structure on the category **Hilb**, and also by restriction on **FHilb**, is defined in the following way:

- **the tensor product** $\otimes\colon \mathbf{Hilb} \times \mathbf{Hilb} \longrightarrow \mathbf{Hilb}$ is the tensor product of Hilbert space;
- **the unit object** I is the one-dimensional Hilbert space $\mathbb{C}$;
- **associators** $(H \otimes J) \otimes K \xrightarrow{\alpha_{H,J,K}} H \otimes (J \otimes K)$ are the unique linear maps satisfying $(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$ for all $u \in H$, $v \in J$ and $w \in K$;
- **left unitors** $\mathbb{C} \otimes H \xrightarrow{\lambda_H} H$ are the unique linear maps satisfying $1 \otimes u \mapsto u$ for all $u \in H$;
- **right unitors** $H \otimes \mathbb{C} \xrightarrow{\rho_H} H$ are the unique linear maps satisfying $u \otimes 1 \mapsto u$ for all $u \in H$.

Although we call the functor $\otimes$ of a monoidal category the 'tensor product', that does not mean that we have to choose the *actual* tensor product of Hilbert spaces for our monoidal structure. There are other monoidal structures on the category that we could choose; a good example is the direct sum of Hilbert spaces. However, the tensor product we have defined above has a special status, since it correctly describes the state space of a composite system in quantum theory.

While **Hilb** is relevant for quantum computation, the monoidal category **Set** is an important setting for classical computation.

Definition 1.4. The monoidal structure on the category **Set**, and also by restriction on **FSet**, is defined as follows for all $a \in A$, $b \in B$ and $c \in C$:

- **the tensor product** is Cartesian product of sets, written $\times$, acting on functions $A \xrightarrow{f} B$ and $C \xrightarrow{g} D$ as $(f \times g)(a, c) = \big(f(a), g(c)\big)$;
- **the unit object** is a chosen singleton set $\{\bullet\}$;
- **associators** $(A \times B) \times C \xrightarrow{\alpha_{A,B,C}} A \times (B \times C)$ are the functions given by $\big((a, b), c\big) \mapsto \big(a, (b, c)\big)$;
- **left unitors** $I \times A \xrightarrow{\lambda_A} A$ are the functions $(\bullet, a) \mapsto a$;
- **right unitors** $A \times I \xrightarrow{\rho_A} A$ are the functions $(a, \bullet) \mapsto a$.

The Cartesian product in **Set** is a categorical product. This is an example of a general phenomenon: if a category has products, then these can be used to give a monoidal structure on the category. The same is true for coproducts, which in **Set** are given by disjoint union.

This highlights an important difference between the standard tensor products on **Hilb** and **Set**: while the tensor product on **Set** comes from a categorical product, the tensor product on **Hilb** does not. We will discover many more differences between **Hilb** and **Set**, which provide insight into the differences between quantum and classical information.

There is a canonical monoidal structure on the category **Rel**.

Definition 1.5. The monoidal structure on the category **Rel** is defined in the following way, for all $a \in A$, $b \in B$, $c \in C$ and $d \in D$:

- **the tensor product** is Cartesian product of sets, written $\times$, acting on relations $A \xrightarrow{R} B$ and $C \xrightarrow{S} D$ by setting $(a, c)(R \times S)(b, d)$ if and only if aRb and cSd;
- **the unit object** is a chosen singleton set $= \{\bullet\}$;
- **associators** $(A \times B) \times C \xrightarrow{\alpha_{A,B,C}} A \times (B \times C)$ are the relations defined by $\big((a, b), c\big) \sim \big(a, (b, c)\big)$;
- **left unitors** $I \times A \xrightarrow{\lambda_A} A$ are the relations defined by $(\bullet, a) \sim a$;
- **right unitors** $A \times I \xrightarrow{\rho_A} A$ are the relations defined by $(a, \bullet) \sim a$.

The Cartesian product is *not* a categorical product in **Rel**, so although this monoidal structure looks like that of **Set**, it is in fact more similar to the structure on **Hilb**.

Example 1.6. If $\mathbf{C}$ is a monoidal category, then so is its opposite $\mathbf{C}^{\mathrm{op}}$. The tensor unit I in $\mathbf{C}^{\mathrm{op}}$ is the same as that in $\mathbf{C}$, whereas the tensor product $A \otimes B$ in $\mathbf{C}^{\mathrm{op}}$ is given by $B \otimes A$ in $\mathbf{C}$, the associators in $\mathbf{C}^{\mathrm{op}}$ are the inverses of those morphisms in $\mathbf{C}$, and the left and right unitors of $\mathbf{C}$ swap roles in $\mathbf{C}^{\mathrm{op}}$.

Monoidal categories have an important property called the *interchange law*, which governs the interaction between the categorical composition and tensor product.

Theorem 1.7 (Interchange)**.** *Any morphisms $A \xrightarrow{f} B$, $B \xrightarrow{g} C$, $D \xrightarrow{h} E$ and $E \xrightarrow{j} F$ in a monoidal category satisfy the interchange law:*

$$(g \circ f) \otimes (j \circ h) = (g \otimes j) \circ (f \otimes h) \tag{1.4}$$

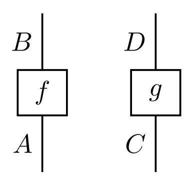
Proof. This holds because of properties of the category $\mathbf{C} \times \mathbf{C}$, and from the fact that $\otimes \colon \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ is a functor:

$$\begin{aligned}
(g \circ f) \otimes (j \circ h) &\equiv \otimes(g \circ f, j \circ h) \\
&= \otimes\big((g, j) \circ (f, h)\big) && \text{(composition in } \mathbf{C} \times \mathbf{C}\text{)} \\
&= \big(\otimes(g, j)\big) \circ \big(\otimes(f, h)\big) && \text{(functoriality of } \otimes\text{)} \\
&= (g \otimes j) \circ (f \otimes h)
\end{aligned}$$

Recall that the functoriality property for a functor F says that $F(g \circ f) = F(g) \circ F(f)$. $\square$

1.2 Graphical calculus

A monoidal structure allows us to interpret multiple processes in our category taking place at the same time. For morphisms $A \xrightarrow{f} B$ and $C \xrightarrow{g} D$, it therefore seems reasonable, at least informally, to draw their tensor product $A \otimes C \xrightarrow{f \otimes g} B \otimes D$ like this:



(1.5)

The idea is that f and g represent processes taking place at the same time on distinct systems. Inputs are drawn at the bottom, and outputs are drawn at the top; in this sense, "time" runs upwards. This extends the one-dimensional notation for categories. Whereas the graphical calculus for ordinary categories was one-dimensional, or *linear*, the graphical calculus for monoidal categories is two-dimensional or *planar*. The two dimensions correspond to the two ways to combine morphisms: by categorical composition (vertically) or by tensor product (horizontally).

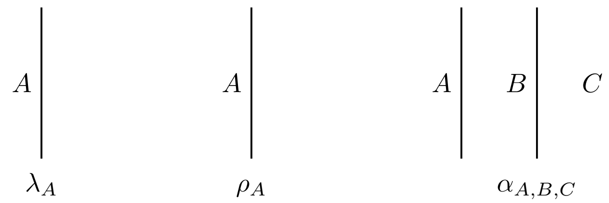
One could imagine this notation being a useful short-hand when working with monoidal categories. This is true, but in fact a lot more can be said: the graphical calculus gives a sound and complete language for monoidal categories.

The (identity on the) monoidal unit object I is drawn as the empty diagram:

(1.6)

The left unitor $I \otimes A \xrightarrow{\lambda_A} A$, the right unitor $A \otimes I \xrightarrow{\rho_A} A$ and the associator $(A \otimes B) \otimes C \xrightarrow{\alpha_{A,B,C}} A \otimes (B \otimes C)$

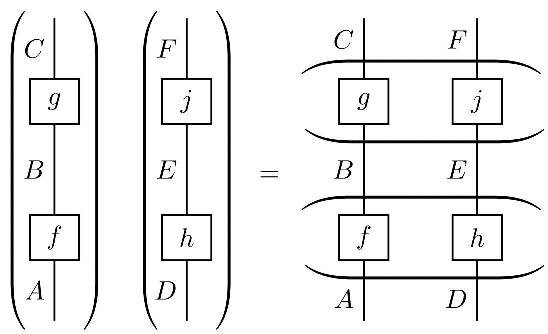
are also simply not depicted:



(1.7)

The coherence of α, λ and ρ is therefore important for the graphical calculus to function: since there can only be a single morphism built from their components of any given type (see Section **??**), it doesn't matter that their graphical calculus encodes no information.

Now consider the graphical representation of the interchange law (1.4):



(1.8)

We use brackets to indicate how we are forming the diagrams on each side. Dropping the brackets, we see the interchange law is very natural; what seemed to be a mysterious algebraic identity becomes clear from the graphical perspective.

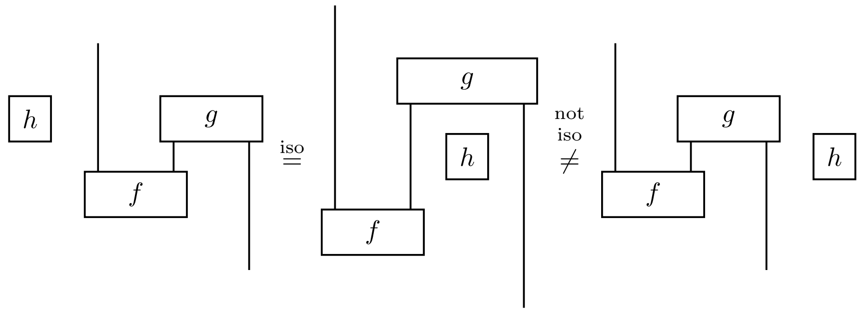
The point of the graphical calculus is that all the superficially complex aspects of the algebraic definition of monoidal categories—the unit law, the associativity law, associators, left unitors, right unitors, the triangle equation, the pentagon equation, the interchange law—melt away, allowing us to make use of the theory of monoidal categories in a direct way. These algebraic features are still there, but they are absorbed into the geometry of the plane, of which our species has an excellent intuitive understanding.

The following theorem is the formal statement that connects the graphical calculus to the theory of monoidal categories.

Theorem 1.8 (Correctness of the graphical calculus for monoidal categories)**.** *A well-typed equation between morphisms in a monoidal category follows from the axioms if and only if it holds in the graphical language up to planar isotopy.*

Two diagrams are *planar isotopic* when one can be deformed continuously into the other within some rectangular region of the plane, with the input and output wires terminating at the lower and upper boundaries of the rectangle, without introducing any intersections of the components. For this purpose, we assume that wires have zero width, and morphism boxes have zero size.

Example 1.9. Here are examples of isotopic and non-isotopic diagrams:



As we have done here, we will often allow the heights of the diagrams to change, and allow input and output wires to slide horizontally along their respective boundaries, although they must never change order. The third diagram here is not isotopic to the first two, since for the h box to move to the right-hand side, it would have to "pass through" one of the wires, which is not allowed. The box cannot pass "over" or "under" the wire, since the diagrams are confined to the plane—that is what is meant by *planar* isotopy. You should imagine that the components of the diagram are trapped between two pieces of glass.

The correctness theorem is really saying two distinct things: that the graphical calculus is *sound*, and that it is *complete*. To understand these concepts, let f and g be morphisms such that the equation $f = g$ is well-typed, and consider the following statements:

- $P(f, g)$: 'under the axioms of a monoidal category, $f = g$';
- $Q(f, g)$: 'the graphical representations of f and g are planar isotopic'.

Soundness is the assertion that for all such f and g, $P(f, g) \Rightarrow Q(f, g)$. *Completeness* is the reverse assertion, that $Q(f, g) \Rightarrow P(f, g)$ for all such f and g.

Proving soundness is straightforward: there are only a finite number of axioms, and one just has to check that they are all valid in terms of planar isotopy of diagrams. Completeness is much harder, and beyond the scope of this book: one must analyze the definition of planar isotopy, and show that any planar isotopy can be built from a small set of moves, each of which independently leave the value of the morphism in the monoidal category unchanged.

Let's take a closer look at the condition that the equation $f = g$ must be well-typed. Firstly, f and g must have the same source and the same target. For example, let $f = \mathrm{id}_{A\otimes B}$, and $g = \rho_A \otimes \mathrm{id}_B$. Then their types are $A \otimes B \xrightarrow{f} A \otimes B$ and $(A \otimes I) \otimes B \xrightarrow{g} A \otimes B$. These have different source objects, and so the equation is not well-typed, even though their graphical representations are planar isotopic. Also, suppose that our category happened to satisfy $A \otimes B = (A \otimes I) \otimes B$; then although f and g would have the same type, the equation $f = g$ would still not be well-typed, since it would be making use of this 'accidental' equality. For a careful examination of the well-typed property.

The notation $\overset{\text{iso}}{=}$ to denote isotopic diagrams, whose interpretations as morphisms in a monoidal category are therefore equal, will be used throughout this book, to indicate an application of the correctness property of the graphical calculus.

1.3 States and effects

If a mathematical structure lives as an object of a category, and we want to learn something about its internal structure, we must find a way to do it using the morphisms of the category only. For example, consider a set $A \in \mathrm{Ob}(\mathbf{Set})$ with a chosen element $a \in A$: we can represent this with the function $\{\bullet\} \longrightarrow A$ defined by $\bullet \mapsto a$. This inspires the following definition, which gives us a generalized categorical notion of state.

Definition 1.10. In a monoidal category, a *state* of an object A is a morphism $I \longrightarrow A$. States are sometimes also called *points*.

Since the monoidal unit object represents the trivial system, a state $I \longrightarrow A$ of a system can be thought of as a way for the system A to be brought into being.

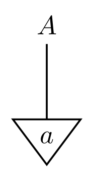
Example 1.11. We now examine what the states are in our three example categories:

- in **Hilb**, states of a Hilbert space H are linear functions $\mathbb{C} \longrightarrow H$, which correspond to elements of H by considering the image of $1 \in \mathbb{C}$;
- in **Set**, states of a set A are functions $\{\bullet\} \longrightarrow A$, which correspond to elements of A by considering the image of $\bullet$;
- in **Rel**, states of a set A are relations $\{\bullet\} \xrightarrow{R} A$, which correspond to subsets of A by considering all elements related to $\bullet$.

Definition 1.12. A monoidal category is *well-pointed* if for all parallel pairs of morphisms $A \xrightarrow{f,g} B$, we have $f = g$ when $f \circ a = g \circ a$ for all states $I \xrightarrow{a} A$. A monoidal category is *monoidally well-pointed* if for all parallel pairs of morphisms $A_1 \otimes \cdots \otimes A_n \xrightarrow{f,g} B$, we have $f = g$ when $f \circ (a_1 \otimes \cdots \otimes a_n) = g \circ (a_1 \otimes \cdots \otimes a_n)$ for all states $I \xrightarrow{a_1} A_1, \ldots, I \xrightarrow{a_n} A_n$.

The idea is that in a well-pointed category, we can tell whether or not morphisms are equal just by seeing how they affect states of their domain objects. In a monoidally well-pointed category, it is even enough to consider product states to verify equality of morphisms out of a compound object. The categories **Set**, **Rel**, **Vect**, and **Hilb** are all monoidally well-pointed. For the latter two, this comes down to the fact that if $\{d_i\}$ is a basis for H and $\{e_j\}$ is a basis for K, then $\{d_i \otimes e_j\}$ is a basis for $H \otimes K$.

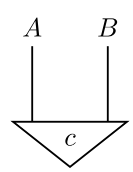
To emphasize that states $I \xrightarrow{a} A$ have the empty picture (1.6) as their domain, we will draw them as triangles instead of boxes.



(1.9)

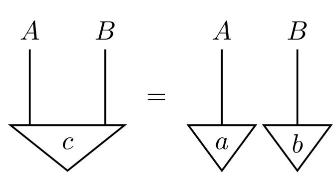
1.4 Product states and entangled states

For objects A and B of a monoidal category, a morphism $I \xrightarrow{c} A \otimes B$ is a *joint state* of A and B. We depict it graphically in the following way.



(1.10)

Definition 1.13. A joint state $I \xrightarrow{c} A \otimes B$ is a *product state* when it is of the form $I \xrightarrow{\lambda_I^{-1}} I \otimes I \xrightarrow{a \otimes b} A \otimes B$ for $I \xrightarrow{a} A$ and $I \xrightarrow{b} B$:



(1.11)

Definition 1.14. A joint state is *entangled* when it is not a product state.

Entangled states represent preparations of $A \otimes B$ which cannot be decomposed as a preparation of A alongside a preparation of B. In this case, there is some essential connection between A and B which means that they cannot have been prepared independently.

Example 1.15. Joint states, product states, and entangled states look as follows in our example categories:

- in **Hilb**:
  - **joint states** of H and K are elements of $H \otimes K$;
  - **product states** are factorizable states;
  - **entangled states** are elements of $H \otimes K$ which cannot be factorized;
- in **Set**:
  - **joint states** of A and B are elements of $A \times B$;
  - **product states** are elements $(a, b) \in A \times B$ coming from $a \in A$ and $b \in B$;
  - **entangled states** don't exist;
- in **Rel**:
  - **joint states** of A and B are subsets of $A \times B$;
  - **product states** are subsets $U \subseteq A \times B$ such that, for some $V \subseteq A$ and $W \subseteq B$, $(v, w) \in U$ if and only if $v \in V$ and $w \in W$;
  - **entangled states** are subsets of $A \times B$ that are not of this form.

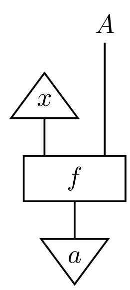
This hints at why entanglement can be difficult to understand intuitively: classically, in the processes encoded by the category **Set**, it cannot occur. However, if we allow nondeterministic behaviour as encoded by **Rel**, then an analogue of entanglement does appear.

1.5 Effects

An *effect* represents a process by which a system is destroyed, or consumed.

Definition 1.16. In a monoidal category, an *effect* or *costate* for an object A is a morphism $A \longrightarrow I$.

Given a diagram constructed using the graphical calculus, we can interpret it as a history of events that have taken place. If the diagram contains an effect, this is interpreted as the assertion that a measurement was performed, with the given effect as the result. For example, an interesting diagram would be this one:



(1.12)

This describes a history in which a state a is prepared, and then a process f is performed producing two systems, the first of which is measured giving outcome x. This does not imply that the effect x was the *only* possible outcome for the measurement; just that by drawing this diagram, we are only interested in the cases when the outcome x does occur. An effect can be thought of as a *postselection*: we run our entire experiment repeatedly, only accepting the result when we find that our measurement had the specified outcome.

Overall our history is a morphism of type $I \longrightarrow A$, which is a state of A. The postselection interpretation tells us how to prepare this state, given the ability to perform its components.

Example 1.17. These statements are at a very general level. To say more, we must take account of the particular theory of processes described by the monoidal category in which we are working.

- In quantum theory, as encoded by **Hilb**, we require a, f and x to be partial isometries. The rules of quantum mechanics then dictate that the probability for this history to take place is given by the square norm of the resulting state. So in particular, the history described by this composite is impossible exactly when the overall state is zero.
- In nondeterministic classical physics, as described by **Rel**, we need put no particular requirements on a, f and x — they may be arbitrary relations of the correct types. The overall composite relation then describes the possible ways in which A can be prepared as a result of this history. If the overall composite is empty, that means this particular sequence of a state preparation, a dynamics step, and a measurement result cannot occur.
- Things are very different in **Set**. The monoidal unit object is *terminal* in that category, meaning $\mathbf{Set}(A, I)$ has only a single element for any object A. So every object has a *unique* effect, and there is no nontrivial notion of 'measurement'.

1.6 Braiding and symmetry

In many theories of processes, if A and B are systems, the systems $A \otimes B$ and $B \otimes A$ can be considered essentially equivalent. While we would not expect them to be equal, we might at least expect there to be some special process of type $A \otimes B \to B \otimes A$ that 'switches' the systems, and does nothing more. Developing these ideas gives rise to *braided* and *symmetric* monoidal categories, which we now investigate.

We first consider braided monoidal categories.

Definition 1.18. A *braided monoidal category* is a monoidal category equipped with a natural isomorphism

$$A \otimes B \xrightarrow{\sigma_{A,B}} B \otimes A \tag{1.13}$$

satisfying the following *hexagon equations*:

$$\begin{tikzcd}
& A \otimes (B \otimes C) \arrow[r, "\sigma_{A,B\otimes C}"] \arrow[dl, "\alpha^{-1}_{A,B,C}"'] & (B \otimes C) \otimes A & \\
(A \otimes B) \otimes C \arrow[dr, "\sigma_{A,B} \otimes \mathrm{id}_C"'] & & & B \otimes (C \otimes A) \arrow[ul, "\alpha^{-1}_{B,C,A}"'] \\
& (B \otimes A) \otimes C \arrow[r, "\alpha_{B,A,C}"'] & B \otimes (A \otimes C) \arrow[ur, "\mathrm{id}_B \otimes \sigma_{A,C}"'] &
\end{tikzcd} \tag{1.14}$$

$$\begin{tikzcd}
& (A \otimes B) \otimes C \arrow[r, "\sigma_{A\otimes B,C}"] \arrow[dl, "\alpha_{A,B,C}"'] & C \otimes (A \otimes B) & \\
A \otimes (B \otimes C) \arrow[dr, "\mathrm{id}_A \otimes \sigma_{B,C}"'] & & & (C \otimes A) \otimes B \arrow[ul, "\alpha_{C,A,B}"'] \\
& A \otimes (C \otimes B) \arrow[r, "\alpha^{-1}_{A,C,B}"'] & (A \otimes C) \otimes B \arrow[ur, "\sigma_{A,C} \otimes \mathrm{id}_B"'] &
\end{tikzcd} \tag{1.15}$$

We include the braiding in the graphical notation like this:

$$A \otimes B \xrightarrow{\sigma_{A,B}} B \otimes A \qquad\qquad B \otimes A \xrightarrow{\sigma^{-1}_{A,B}} A \otimes B \tag{1.16}$$

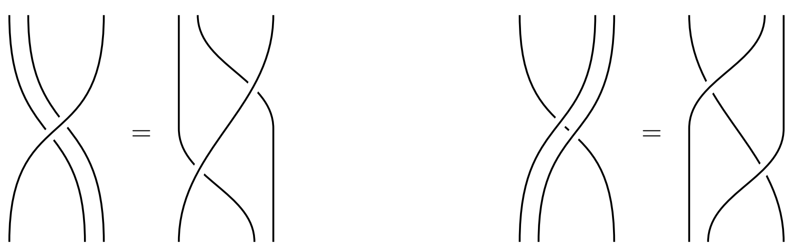
Invertibility then takes the following graphical form:

(1.17)

This captures part of the geometric behaviour of strings. Naturality of the braiding and the inverse braiding have the following graphical representations:

(1.18)

The hexagon equations have the following graphical representations:



(1.19)

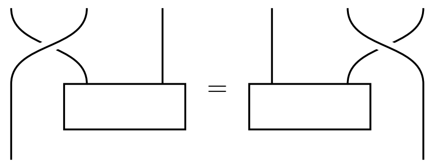
Each of these equations has two strands close to each other on the left-hand side, to indicate that we are treating them as a single composite object for the purposes of the braiding. We see that the hexagon equations are saying something quite straightforward: to braid with a tensor product of two strands is the same as braiding separately with one then the other.

Since the strands of a braiding cross over each other, they are not lying on the plane; they live in three-dimensional space. So while categories have a one-dimensional or linear notation, and monoidal categories have a two-dimensional or planar graphical notation, braided monoidal categories have a three-dimensional notation. Because of this, braided monoidal categories have an important connection to three-dimensional quantum field theory.

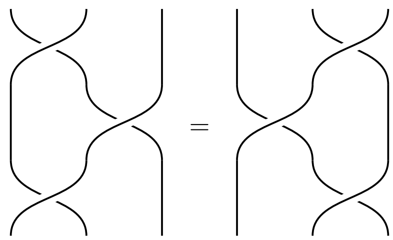
Braided monoidal categories have a sound and complete graphical calculus, as established by the following theorem. The notion of isotopy it uses is now three-dimensional; that is, the diagrams are assumed to lie in a cube, with input wires terminating at the lower face and output wires terminating at the upper face. This is also called *spatial isotopy*.

Theorem 1.19 (Correctness of graphical calculus for braided monoidal categories)**.** *A well-typed equation between morphisms in a braided monoidal category follows from the axioms if and only if it holds in the graphical language up to spatial isotopy.*

Given two isotopic diagrams, it can be quite nontrivial to show they are equal using the axioms of braided monoidal categories directly. So as with ordinary monoidal categories, the coherence theorem is quite powerful. For example, try to show that the following two equations hold directly using the axioms of a braided monoidal category:



(1.20)



(1.21)

Equation (1.21) is called the *Yang–Baxter equation*, which plays an important role in the mathematical theory of knots.

We now give some examples of braided monoidal categories. For each of our main example categories there is a naive notion of a 'swap' process, which in each case gives a braided monoidal structure.

Definition 1.20. Our example categories **Hilb**, **Set** and **Rel** can all be equipped with a canonical braiding:

- in **Hilb**, $H \otimes K \xrightarrow{\sigma_{H,K}} K \otimes H$ is the unique linear map extending $a \otimes b \mapsto b \otimes a$ for all $a \in H$ and $b \in K$;
- in **Set**, $A \times B \xrightarrow{\sigma_{A,B}} B \times A$ is defined by $(a, b) \mapsto (b, a)$ for all $a \in A$ and $b \in B$;
- in **Rel**, $A \times B \xrightarrow{\sigma_{A,B}} B \times A$ is defined by $(a, b) \sim (b, a)$ for all $a \in A$ and $b \in B$.

In fact these are all symmetric monoidal structures, which we explore in Section 1.7.

1.7 Symmetric monoidal categories

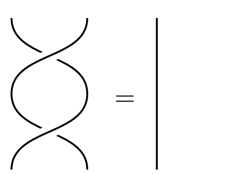
In our example categories **Hilb**, **Rel** and **Set**, the braidings satisfy an extra property that makes them very easy to work with.

Definition 1.21. A braided monoidal category is *symmetric* when

$$\sigma_{B,A} \circ \sigma_{A,B} = \mathrm{id}_{A \otimes B} \tag{1.22}$$

for all objects A and B, in which case we call σ the *symmetry*.

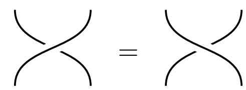
Graphically, condition (1.22) has the following representation.



(1.23)

Intuitively: the strings can pass through each other, and nontrivial knots cannot be formed.

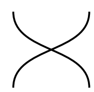
Lemma 1.22. *In a symmetric monoidal category $\sigma_{A,B} = \sigma^{-1}_{B,A}$, with the following graphical representation:*



(1.24)

Proof. Combine (1.17) and (1.23). □

A symmetric monoidal category therefore makes no distinction between over- and under-crossings, and so we simplify our graphical notation, drawing



(1.25)

for the single type of crossing.

The graphical calculus with the extension of braiding or symmetry is still sound: if the two diagrams of morphisms can be deformed into one another, then the two morphisms are equal.

Suppose we imagine our diagrams as curves embedded in four-dimensional space. Then we can smoothly deform one crossing into the other, in the manner of equation (1.24), by making use of the extra dimension. In this sense, symmetric monoidal categories have a four-dimensional graphical notation. The following correctness theorem therefore uses the four-dimensional version of isotopy.

Theorem 1.23 (Correctness of the graphical calculus for symmetric monoidal categories)**.** *A well-typed equation between morphisms in a symmetric monoidal category follows from the axioms if and only if it holds in the graphical language up to four-dimensional isotopy.*