6

Quantum Processes

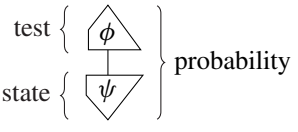
> The art of progress is to preserve order amid change, and to preserve change amid order.
>
> – *Alfred North Whitehead, Process and Reality, 1929*

From now on, no more babies, plug-strips, or cooking: we now focus exclusively on quantum processes. Naturally, our first goal is to construct the theory of **quantum processes**. Fortunately, the work that we have done so far brings us fairly close to that goal.

For one thing, the process theory of **linear maps** is not too far away from full-blown **quantum processes**. In particular, the theory of **quantum processes** will inherit its string diagram description from **linear maps**. This provides us with a high-level language that will make reasoning about quantum processes very easy, for example in the context of quantum computation and quantum protocol design. This also means that we already know several features of quantum processes from Section 4.4, namely those that happen to hold in all process theories admitting string diagrams: the existence of non-separable states, the no-cloning theorem, teleportation, and so on.

Note that we say **linear maps** are not too far away from **quantum processes**, not that they are **quantum processes**. In this chapter, we will proceed from **linear maps** to **quantum processes** in a few steps (Fig. 6.1), which correct the shortcomings of the former. A first issue is that, as a model of quantum processes, **linear maps** contain some redundant data, namely 'global phases'. These will never be detectable by quantum measurements (which we'll meet in the next chapter) and hence have no discernable effect on which process actually happened.

But so what? Who cares? We could just carry on using linear maps, and ignore global phases whenever necessary. In fact, many textbooks on quantum theory would do just that. On the other hand, the generalised Born rule:



(6.1)

wouldn't work, because the numbers generated are not positive real numbers (which we can interpret as probabilities), but complex numbers.

We address this issue by performing a simple construction on the process theory of **linear maps** to produce another theory called **quantum maps**, which truly describes the processes that may take place in quantum theory.

First, in Section 6.1 we turn **linear maps** into **pure quantum maps** by *doubling* every process. This magically solves the problem of global phases and probabilities in one fell swoop.

After this, we show in Section 6.2.1 that pure quantum maps fail to include one very important process: the process of *discarding* a system. Often we wish to ignore some part of a larger system when it is out of our control (e.g. a potential eavesdropper in a security protocol) or simply irrelevant (e.g. some electrons flying around on Mars). By adding this discarding process to pure quantum maps, we obtain **quantum maps**. Many new *impure* (or *mixed*) quantum maps arise by composing pure ones with the discarding process.

An alternative interpretation of the impurity of certain quantum maps is *probabilistic mixing*. In classical physics, probabilities can be seen as a way of accounting for our lack of knowledge about the state of a system. For example a branch of physics called *statistical mechanics* describes the states of a system using probability distributions, since it's nearly impossible to know what each little particle is up to. This is also what probabilistic mixing means in quantum theory: having a lack of knowledge about which process is actually happening.

Note that we said quantum maps describe processes that 'may take place'. Unlike classical physics, quantum theory has processes that are irreducibly non-deterministic. That is, there exist non-deterministic processes that cannot be accounted for solely by a lack of knowledge about the quantum system. Regardless of how perfectly we know the state of a system, such processes will not have a fixed outcome until after they occur. This is the feature of quantum theory that Einstein found deeply upsetting, as he famously said, 'God does not play dice.' To account for this 'quantum dice-throwing', the third and final step in this chapter is to define **quantum processes** as collections of **quantum maps** that together make up the alternatives of what may happen. The rule that tells us which **quantum maps** together make up valid **quantum processes** is called the *causality postulate*.

These irreducibly non-deterministic processes are absolutely essential to quantum theory, for at least two reasons:

1. *Quantum measurement*, which is our only means of interacting with quantum systems, is a non-deterministic quantum process. (Recall Dave's non-deterministic travels narrated in the introductory chapter.)
2. The causality postulate places a strong restriction on which quantum maps can occur as deterministic quantum processes. For example, the only deterministic quantum effect is discarding. On the other hand, every quantum map can be realised as part of a non-deterministic quantum process. This fact is crucial to realising quantum teleportation, among many other things.

The causality postulate is tightly connected to the concept of causality in physics. In particular, we will see in Section 6.3.2 that it forbids 'faster-than-light signalling', hence

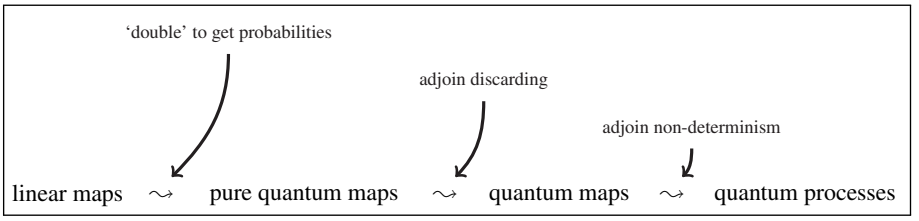


Figure 6.1 The passage from **linear maps** to **quantum processes**.

guaranteeing that quantum theory is not in conflict with that other funky physical theory, *relativity theory*. We don't want to make Einstein even more upset than he already is!

6.1 Pure Quantum Maps from Doubling

In the previous chapter we put a lot work into defining and studying **linear maps**, so one may think that quantum theory is all about these. This is almost true. Recall from Section 3.4.1 that effects can be interpreted as 'tests', and from Section 4.3.1 that the adjoint of a state ψ corresponds to testing whether a system is in the state ψ. When such a test is composed with a state, the number produced should be the probability of that test returning 'yes'.

However, in the case of **linear maps**, if we pick any old states ψ and ϕ, it's quite likely that their inner product won't even be a real number, much less a probability (i.e. a real number between zero and one). Taking this into account, the process theory of **linear maps** is not an appropriate candidate for describing quantum processes. However, we can turn it into one by means of what we call *doubling*, and call the resulting process theory **pure quantum maps**.

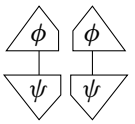
Moreover, doubling has two other nice consequences:

1. It automatically eliminates redundant *global phases* (Section 6.1.2).
2. It makes space for two new ingredients that didn't exist in **linear maps**: the *discarding map* (Section 6.2.1.1) and *classical wires* (Chapter 8).

6.1.1 Doubling Generates Probabilities

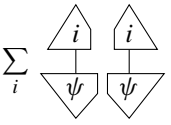
Recall from Proposition 5.70 that if we multiply a number by its conjugate we automatically get a positive number, so:

$$0 \leq \tag{6.2}$$



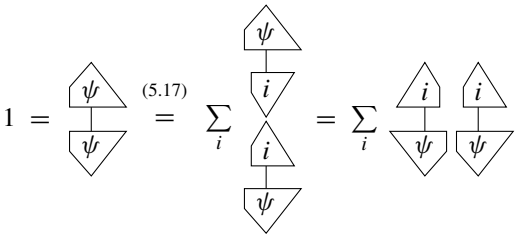
Furthermore, if ψ and ϕ are both normalised, this will be a real number between 0 and 1, i.e. a probability. Rather than proving this fact directly, we prove a more general fact that will be useful later.

Lemma 6.1 For any ONB and any normalised state ψ :



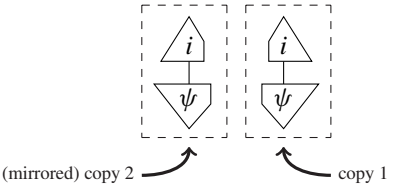
(6.3)

Proof We have:



where the last step uses the fact that numbers are self-transposed. □

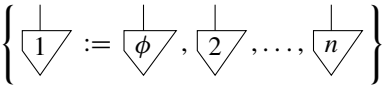
This theorem says that any normalised state along with any ONB, considering 'doubled inner products':



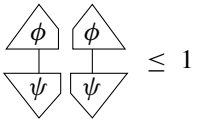
yields a probability distribution (cf. Definition 5.35), that is, a list of positive real numbers that sums up to one.

Remark 6.2 In the next chapter we will see that ONBs represent certain quantum measurements, and then Lemma 6.1 will guarantee that the probabilities for all of the possible outcomes add up to 1.

We showed in Proposition 5.79 that any orthonormal set extends to an ONB. In particular, a single normalised state ϕ extends to an ONB:

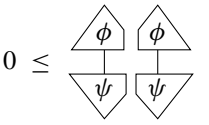


For this ONB, the only way (6.3) can hold is if:



so, we get the following.

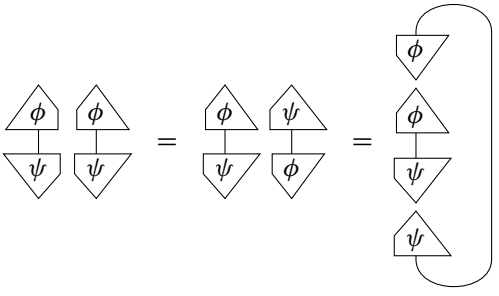
Corollary 6.3 For normalised states and effects we have:



(6.4)

These doubled inner-product diagrams constitute the main mechanism for computing probabilities in quantum theory, and they are what is called the 'Born rule' in standard textbooks. At first, this new thing doesn't look like the generalised Born rule we met back in Section 3.4. However, it will soon!

Remark 6.4 One typically encounters the second and/or the third of the following three equivalent forms of the Born rule:



The first equation follows from the fact that numbers are self-transposed, and the second is just a diagram deformation. In more traditional notation, these alternative forms become, respectively:

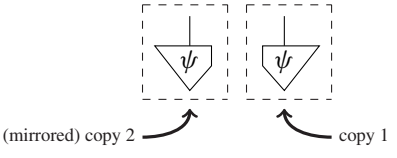
$$\langle\psi|\phi\rangle\langle\phi|\psi\rangle = |\langle\psi|\phi\rangle|^2 \qquad\qquad \mathrm{tr}\big(P_{|\phi\rangle}\,\rho_{|\psi\rangle}\big)$$

where

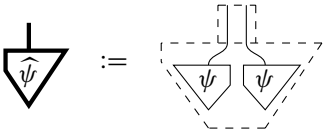
$$P_{|\psi\rangle} = \rho_{|\psi\rangle} := |\psi\rangle\langle\psi|$$

The upshot of our expression of the Born rule is what much of the remainder of this section is about: by transforming states and effects into a 'doubled' form, the Born rule reduces to the simplest way we can produce a number from a state and an effect, namely composing them.

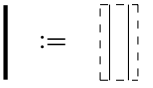
A key point is that the probabilities in the doubled inner product (6.4) do not depend precisely on ψ but rather on 'ψ-doubled':



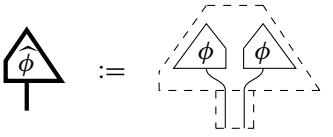
So in order to realise them as an instance of the generalised Born rule (6.1), we can simply treat ψ-doubled as a first-class citizen. That is, we treat it as a state:



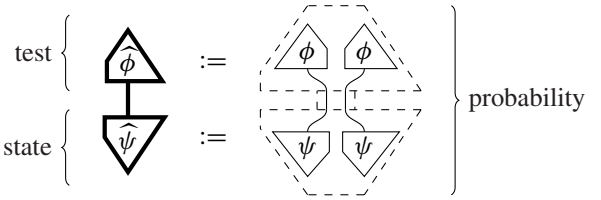
in a new 'doubled-process theory'. In other words, for any state ψ of a Hilbert space A we define a new state $\widehat{\psi}$ that is the doubled version of ψ, and to $\widehat{\psi}$ we attribute a new type $\widehat{A}$, which is secretly just two copies of A. Diagrammatically:



Similarly, we define new effects for this new type $\widehat{A}$:

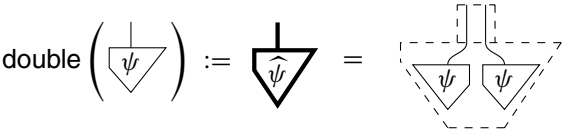


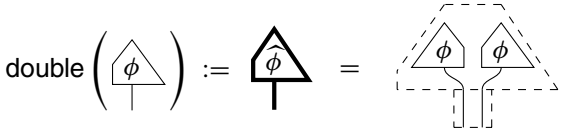
Together the new state and effect yield:



Bingo! We now see the Born rule (6.4)) from quantum theory arising as a special case of the generalised Born rule (6.1).

We call these doubled states and effects *pure quantum states* and *pure quantum effects*, respectively. We sometimes will use the following notation:

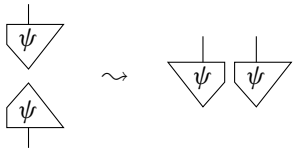




Remark 6.5 This 'doubling trick' is closely related to a construction that is familiar in quantum theory, namely, the passage from a pure state vector $|\psi\rangle$ to its associated *density operator*:

$$\widetilde{\psi} := |\psi\rangle\langle\psi|$$

This has the same data as a doubled state, which can be seen just by transposing the effect ψ into the conjugate of the state ψ:



If ψ is normalised, then by (4.53) the density operator $\widetilde{\psi}$ is a projector. These projectors play the role of quantum states in the traditional literature. However, the process-theoretic paradigm takes states to be processes with no inputs. Clearly, density operators break this convention, which becomes a bit of a pain later on (cf. Remark 6.50). On the other hand, our doubled states (conveniently) retain it.

6.1.2 Doubling Eliminates Global Phases

As we already mentioned, the probabilities produced by the Born rule do not depend precisely on ψ, but rather on 'ψ-doubled'. This distinction might seem trivial, until one realises that the correspondence between states and doubled states is not one-to-one, but rather many-to-one. This phenomenon already occurs for numbers. There are many complex numbers λ such that:

$$\overline{\lambda}\lambda = 1 \tag{6.5}$$

for example, $1, -1, i, -i$.

In Section 5.3.1, we saw that it is possible to write any complex number as:

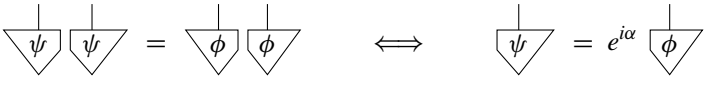
$$re^{i\alpha}$$

Equation (6.5) just means $r = 1$, so any such λ is a number of the form:

$$e^{i\alpha}$$

By definition, these numbers vanish when they are multiplied by their conjugate, so they have no effect on a doubled state. In fact, this is the only data that gets lost in the doubled state.

Proposition 6.6 Two states ψ and ϕ become the same state when doubled if and only if they are equal up to some number $e^{i\alpha}$, i.e.:



for $\alpha \in [0, 2\pi)$. The number $e^{i\alpha}$ is called a *global phase*.

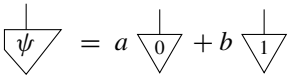
The proof is an instance of a more general one, given below as the proof of Theorem 6.17.

In fact, from the very start of quantum theory, as formulated by von Neumann, global phases were declared meaningless but remained an explicit part of the formalism. Most textbooks on quantum theory deal with this fact by reserving the term 'quantum state' for an 'equivalence class' of states, namely those that are equal up to a global phase. However, doubling gives us a more elegant and simpler way to deal with this problem.

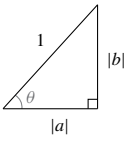
The usual justification for declaring global phases physically meaningless is that they are not empirically accessible; i.e. they cannot be discovered by means of quantum measurements, which we'll study in Chapter 7. This is already apparent from the fact that all of the probabilities produced by quantum theory come from the Born rule, which only makes use of doubled states and effects. Thus there is really no point in distinguishing two states that differ only by a global phase.

Ignoring global phases has a useful practical consequence as well: it allows our puny human brains to actually picture quantum systems in a geometric way. This is something quite handy for physics, and we do it all the time.

The states of the simplest non-trivial quantum system, *qubits*, live in $\widehat{\mathbb{C}^2}$. Since we can represent such a state with two complex numbers, we can do it with four real numbers. So, naïvely, one may think we need four-dimensional space to write down the state of a qubit. But that's one too many dimensions for (most) humans to picture! The job becomes much easier by just looking at normalised states. Suppose:

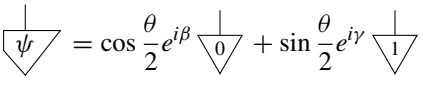


Then, if ψ is normalised, it must be the case that $|a|^2 + |b|^2 = 1$. Thus, if we want to know $|a|$ and $|b|$, we can ask Pythagoras:



If you remember your trigonometry, this means $|a| = \cos\theta$ and $|b| = \sin\theta$. If you flunked trig, just take our word for it.

As a matter of convention, it's slightly more convenient to use $\frac{\theta}{2}$ instead of θ, so let $|a| = \cos\frac{\theta}{2}$ and $|b| = \sin\frac{\theta}{2}$. Then, we can then drop the absolute values by introducing complex phases, which gives us:

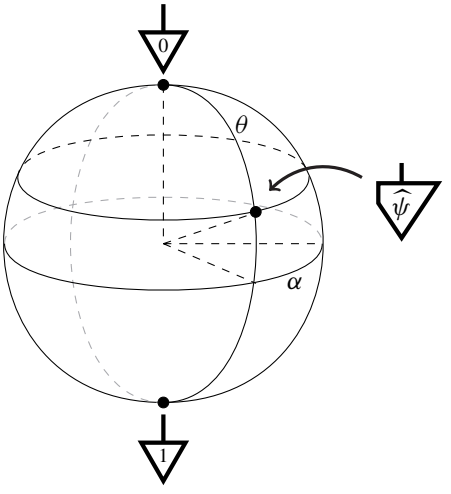


So, we've replaced four real parameters with three angles. This is where doubling comes in. Since doubling kills global phases, the angle β is actually redundant, because we can

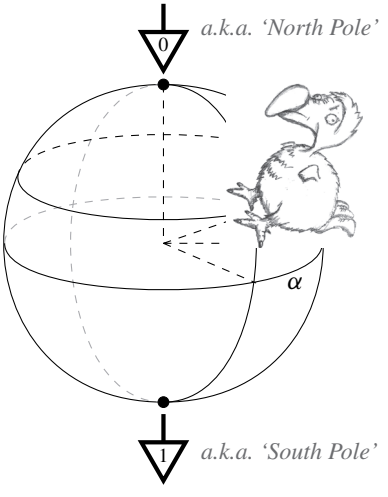
multiply the whole thing by $e^{-i\beta}$. For some α (namely $\alpha := \gamma - \beta$), we can therefore write the quantum state $\widehat{\psi}$ conveniently as:

$$\widehat{\psi} := \text{double}\left(\cos\frac{\theta}{2}\,0 + \sin\frac{\theta}{2}e^{i\alpha}\,1\right)$$

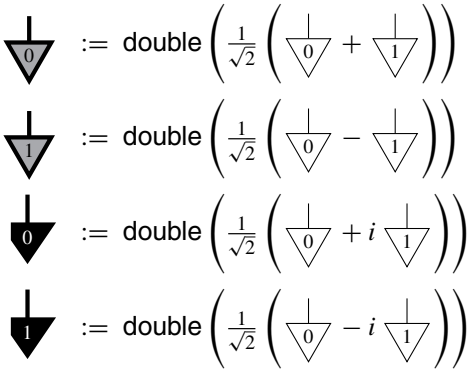
Since the quantum state is now totally described by two angles, we can plot it on a sphere, called the *Bloch sphere*:



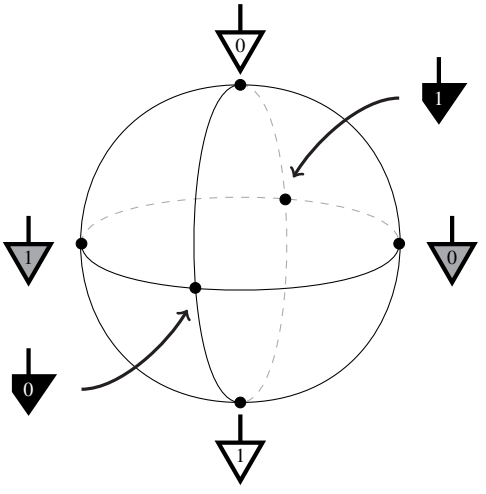
This picture is useful for our intuition. For example, the more 'similar' two states are, that is, the higher the value of their inner product, the closer they are on the Bloch sphere. In particular, orthogonal states are always antipodes. Remember Dave's travels from Section 1.1? We now know what sort of sphere he was hopping around on:



Exercise 6.7 Show that the following points:

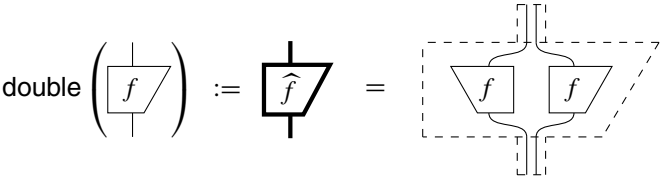


are located on the Bloch sphere as follows:



6.1.3 The Process Theory of Pure Quantum Maps

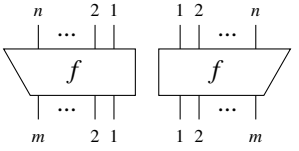
We will now construct a full-blown theory of doubled processes. Besides states and effects, arbitrary linear maps can also be doubled:



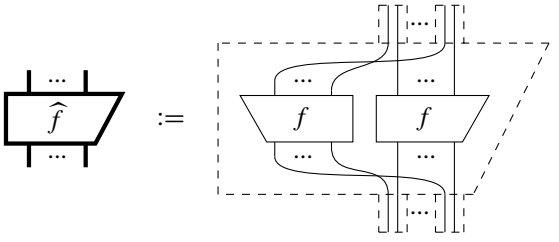
(6.6)

In fact, this extends to processes with any number of inputs and outputs. However, we should be a bit careful with which pairs of thin wires should be taken together to form a

thick wire. We should always pair the first input/output of f with the first input/output of f's conjugate, the second with the second, and so on:



Note that since f's conjugate is the mirror image of f, we count inputs and outputs from right to left rather than left to right. As a result, pairing up inputs and outputs introduces a 'twist' in the wires connected to the conjugate process:



(6.7)

One of the benefits of using the new thick boxes and wires is that the extra complexity of all this twisting remains hidden within the notation.

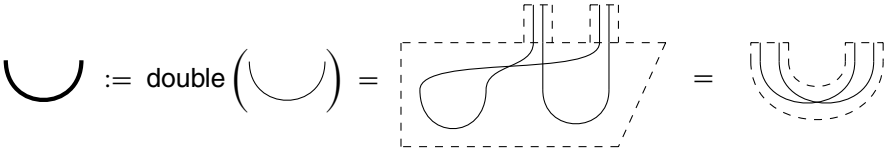
Doubling all processes yields the following new process theory.

Definition 6.8 The process theory of **pure quantum maps** has as types $\widehat{A}$ for all Hilbert spaces A in **linear maps**, and has as processes doubled linear maps $\widehat{f}$ for all processes f in **linear maps**.

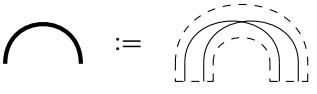
The processes in **pure quantum maps** aren't really all that new; they are just special kinds of linear maps. In other words, **pure quantum maps** is a subtheory of **linear maps**:

$$\textbf{pure quantum maps} \subseteq \textbf{linear maps}$$

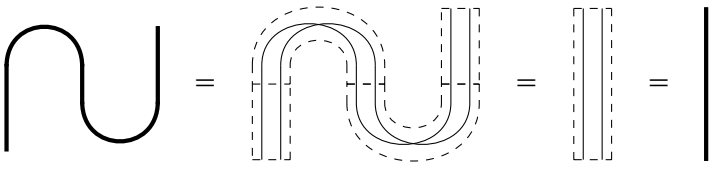
Notably, the fact that **linear maps** admits string diagrams is inherited by **pure quantum maps**. Applying equation (6.7) to the cup yields a 'twisted' double cup:



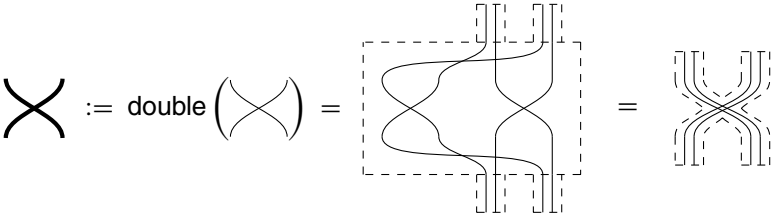
and a corresponding double cap:



When we compose these two, the 'twists' cancel out, yielding the first yanking equation:

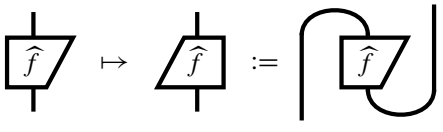


Concerning the other two yanking equations, first applying equation (6.7) to the swap map yields a double swap:



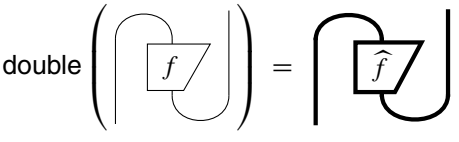
and when we compose it with a cap or a cup, we obtain:

Since there are cups and caps, there also is a notion of transposition for the doubled theory:

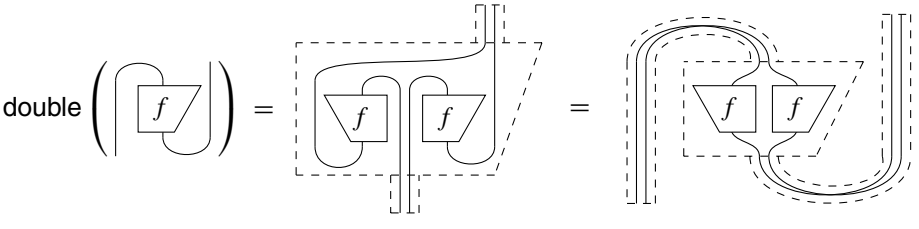


which turns out to coincide with transposition in the undoubled theory.

Proposition 6.9 Doubling preserves transposition:

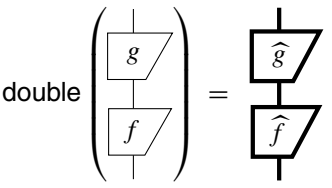


Proof In the single input/output case we have:

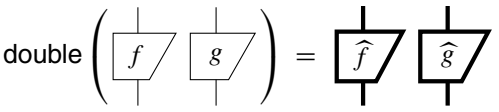


The many input/output case from equation (6.7) can be shown similarly. □

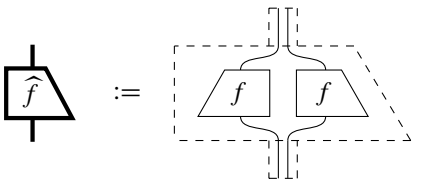
The fact that doubling preserves transposes is an instance of a more general fact that doubling preserves diagrams. This can be best seen by decomposing a string diagram into its constituent processes. We already know that doubling sends cups/caps to cups/caps. Evidently, doubling preserves sequential composition:



Exercise 6.10 Show that doubling preserves parallel composition:



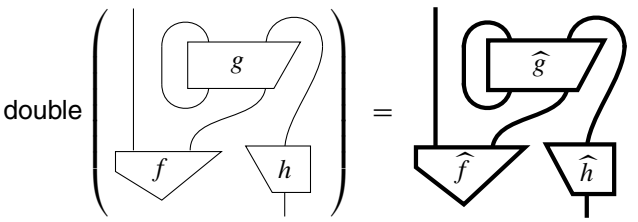
The doubled theory also inherits its adjoints from the non-doubled theory:



(6.8)

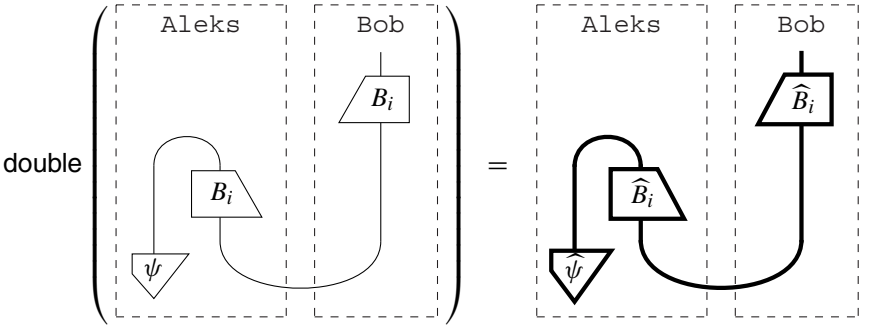
If we put all of these pieces together, we can conclude the following.

Corollary 6.11 Doubling preserves string diagrams:

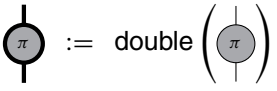


As a result, any of the calculations we have previously done for diagrams of **linear maps** lifts straightforwardly to **pure quantum maps** just by doubling all of the diagrams.

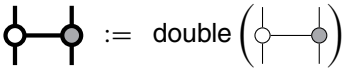
Example 6.12 In Section 5.3.6, we realised teleportation in the theory of **linear maps** using the Bell maps. To pass to quantum maps, we simply double everything (except Aleks and Bob of course):



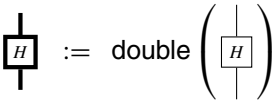
Example 6.13 In Sections 5.3.4 and 5.3.5 we showed how to turn classical logic gates into linear maps. Now, by relying on doubling, we can turn them into *quantum (logic) gates*. For example, the *quantum NOT gate* is:



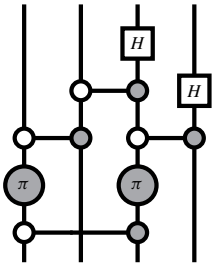
the *quantum CNOT gate* is:



and the *Hadamard gate* is:

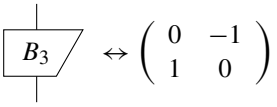


which doesn't even have a classical counterpart. *Quantum circuits* constitute the application of quantum gates to a fixed number of qubits, for example:

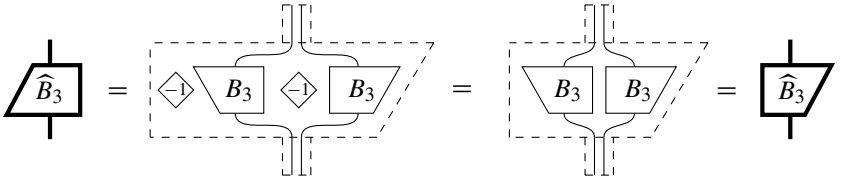


Aside from the Hadamards, we could already do everything above using classical logic gates. Things will start to get much more interesting when we introduce *phase gates* in Chapter 9.

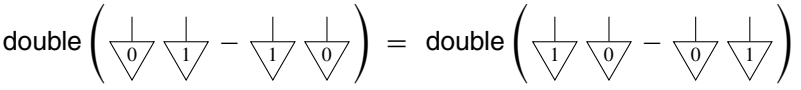
Example 6.14 When doubling the Bell matrices of Section 5.3.6, then:



becomes self-transposed, since doubling eliminates global phases:



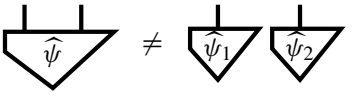
For the corresponding Bell state we then also have:



Generally, we will carry on using the same terminology for **pure quantum maps** that we use for processes in any other theory. However, certain concepts are important enough to get dedicated 'quantum' names.

Definition 6.15 A $\otimes$-non-separable pure quantum state is called an *entangled state*.

So, an entangled pure state is any bipartite state $\widehat{\psi}$ in **pure quantum maps** that does not factor into single-system states $\widehat{\psi}_1$ and $\widehat{\psi}_2$:



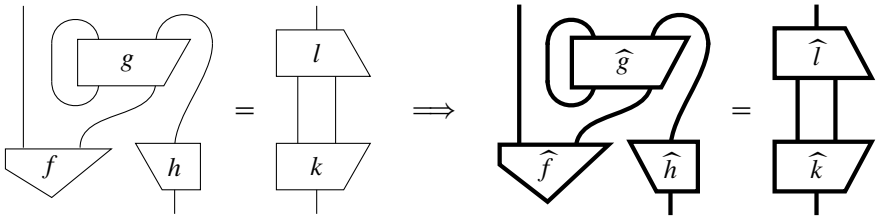
An example is the (doubled) Bell state:

As mentioned in Remark 4.1, we will need to refine this definition of entanglement once we pass from pure states to more general quantum states. We do this in Section 8.3.5.

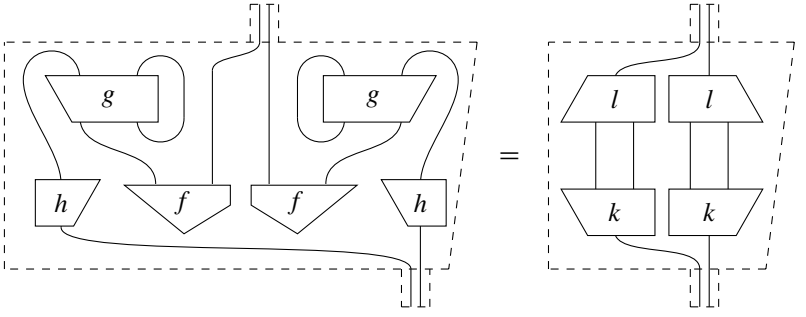
6.1.4 Things Preserved by Doubling

By Corollary 6.11, doubling preserves string diagrams, so if two diagrams are equal, then their doubled counterparts will also be equal.

Corollary 6.16 Any equation that holds between string diagrams of 'single' processes also holds for its doubled version:



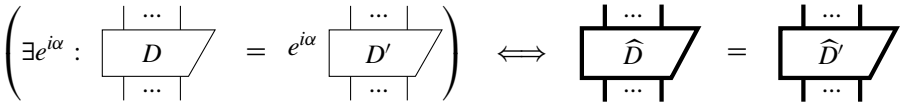
We can see this directly by unfolding both sides of the doubled equation:



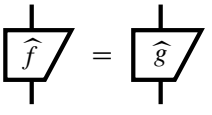
Then clearly the equality between diagrams of 'single' processes together with its conjugated version yields the doubled equation.

The converse of Corollary 6.16 is almost true. Since doubling eliminates global phases, we need to reintroduce them to get 'if and only if'.

Theorem 6.17 Let D and D' be arbitrary diagrams in **linear maps**, and $\widehat{D}$ and $\widehat{D'}$ be their doubled versions in **pure quantum maps**; then:

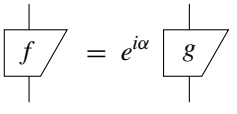


Proof (⇒) directly follows from Corollary 6.16. For (⇐), we can use Corollary 6.11 to replace a diagram of doubled maps with one big, doubled map. So, it suffices to show that for any linear maps f and g such that:



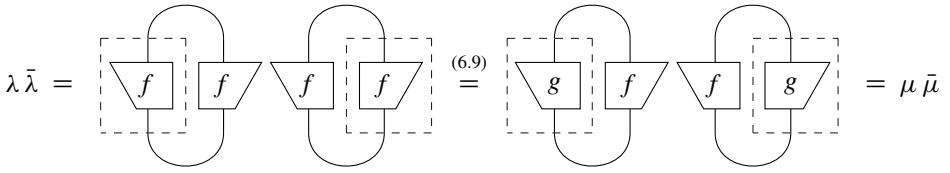
(6.9)

there exists some α such that:



Let λ and μ be defined as:

Then:



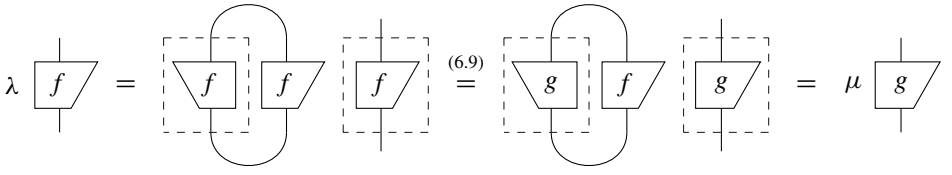
where the dotted lines indicate where we use (6.9). There are two cases: either $\lambda \neq 0$ or $\lambda = 0$. In the first case, divide both sides of the equation above by $\lambda\overline{\lambda}$:

$$1 = \frac{\mu\,\overline{\mu}}{\lambda\,\overline{\lambda}} = \left(\frac{\mu}{\lambda}\right)\overline{\left(\frac{\mu}{\lambda}\right)}$$

So $\frac{\mu}{\lambda}$ is a global phase, i.e.:

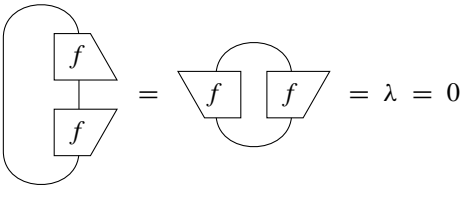
$$\frac{\mu}{\lambda} = e^{i\alpha}$$

for some α, and:



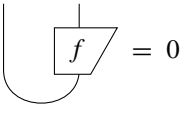
so we indeed obtain $f = e^{i\alpha}\, g$.

In the second case:

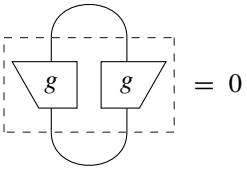


(6.10)

Thus by positive definiteness (see Section 5.3.2):



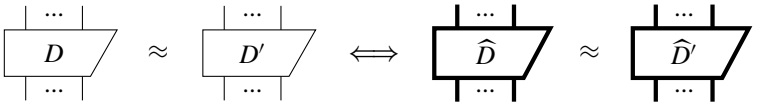
so $f = 0$, hence $\widehat{f} = 0$, and hence by assumption, $\widehat{g} = 0$. This can only be the case if g itself is zero. To see this, attach cups/caps to $\widehat{g}$:



and using positive definiteness just as we did in (6.10) we obtain $g = 0$. Hence $f = e^{i\alpha} g = 0$ where now α can be any angle. □

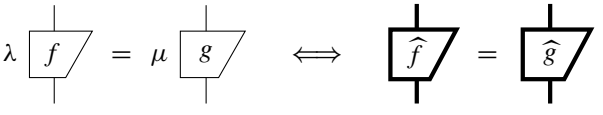
Of course, if in Theorem 6.17 we replace equality with equality up to a number (cf. Section 3.4.3), the phase vanishes.

Corollary 6.18 For D and D' arbitrary diagrams in **linear maps**:



Remark 6.19 Besides what is actually being proven in Theorem 6.17, there is another surprising lesson to be learned from this proof. The crux of this proof is purely diagrammatic! In particular, we define the numbers λ and μ as diagrams:

and we show by means of diagram substitutions that:

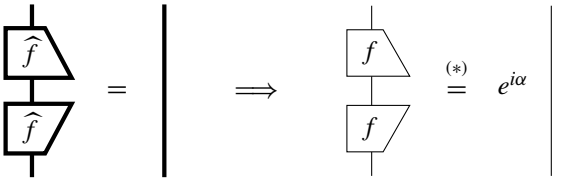


where $\lambda\bar{\lambda} = \mu\bar{\mu}$, and the only place we use some specific structure of **linear maps**, namely the complex numbers, is to put this equation in a form involving $e^{i\alpha}$, which is also what requires the case-distinction. But this can be entirely avoided (see references in Section 6.7). This implies that not just for **linear maps**, but also for any other process theory, doubling identifies precisely those processes that differ only by these 'generalised global phases'. The lesson to take from this is to not judge a book by its cover: what at first looks not at all diagrammatic (e.g. eliminating global phases) may turn out to be fundamentally diagrammatic.

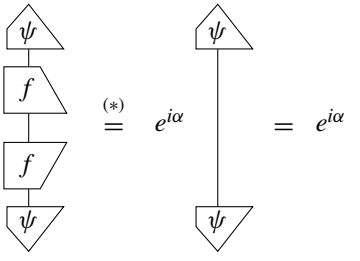
We now show how a number of properties of processes coincide in the doubled world and the undoubled world:

Proposition 6.20 A pure quantum map $\widehat{f}$ is an isometry (respectively unitary) if and only if f is an isometry (respectively unitary).

Proof If f is an isometry, $\widehat{f}$ is an isometry by Corollary 6.16. Conversely, we can use Theorem 6.17 to convert doubled equations into single equations up to a phase:



That is, for $\widehat{f}$ an isometry, f is an isometry up to a global phase $e^{i\alpha}$. We can furthermore show that $\alpha = 0$. First, for any normalised state ψ, we have:

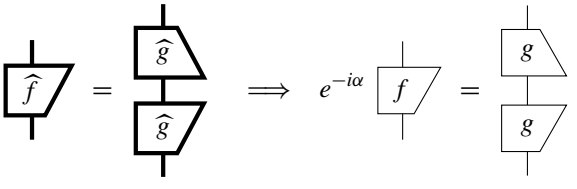


The LHS of the above equation is the inner product of a state with itself, so it must be positive. Since the only global phase that is also a positive real number is 1, $e^{i\alpha} = 1$, so f is an isometry. Unitarity is proven similarly. □

We now prove a result similar to Proposition 6.20 for positive maps and projectors, but with a small caveat. If f is an isometry/unitary, then $e^{i\alpha}f$ is also an isometry/unitary, so we can choose any 'undoubled representative' of $\widehat{f}$. However, this is not the case for positive maps and projectors, where we need to choose a particular representative to get the desired property.

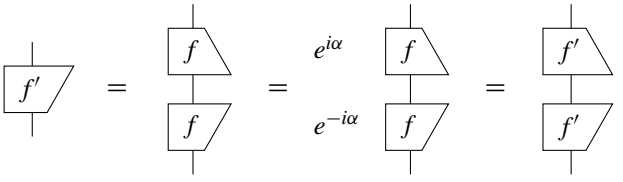
Proposition 6.21 A pure quantum map $\widehat{f}$ is positive (respectively a projector) if and only if there exists a positive linear map (respectively a projector) f' with $\widehat{f'} = \widehat{f}$.

Proof Again, one direction follows from Corollary 6.16. For the other direction, suppose $\widehat{f}$ is positive and apply Theorem 6.17 (moving the global phase to the LHS):



(6.11)

From this, we see that $f' := e^{-i\alpha} f$ is positive. Since doubling removes global phases, $\widehat{f} = \widehat{f'}$. If $\widehat{f}$ is a projector, then (6.11) holds for $g := f$ (cf. Proposition 4.70). Hence f' is also a projector:



□

The following is one more variation on the same theme.

Exercise 6.22 Show that a pure quantum state $\widehat{\psi}$ is normalised if and only if ψ is normalised and that two pure quantum states $\widehat{\psi}$ and $\widehat{\phi}$ are orthogonal if and only if ψ and ϕ are orthogonal.

A consequence of all of the results above is that we can (mostly) work with pure quantum maps as if we were working with plain old linear maps, but with the added benefit that the Born rule is now nothing but a state-effect encounter and that the redundant global phases are eliminated.

Of course there are some exceptions; otherwise, there wouldn't have been any point in doubling everything in the first place.

6.1.5 Things Not Preserved by Doubling

The definition of **linear maps** included three requirements:

- the numbers are the complex numbers;
- there exist sums for processes; and
- there exists an ONB for each type.

Now, we will see that none of these defining features of **linear maps** is preserved under doubling! More specifically:

- the doubled numbers are not the complex numbers;
- sums in the doubled theory are not doubled sums; and
- doubled ONBs are not ONBs in the doubled theory.

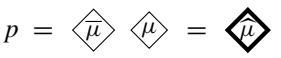
However, the things we obtain instead all play a key role in quantum theory.

6.1.5.1 The Doubled Numbers Are 'Probabilities'

This was of course our initial motivation to do doubling.

Proposition 6.23 The numbers in **pure quantum maps** are the positive real numbers.

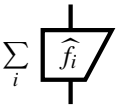
Proof This follows from Proposition 5.70 characterising positive numbers:



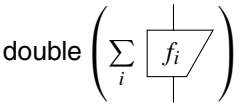
In particular, every positive number is a pure quantum map. □

6.1.5.2 Sums in the Doubled Theory Represent 'Mixing'

Sums of doubled processes:

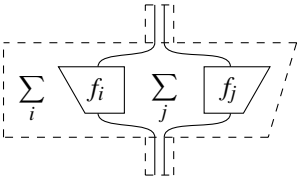


are in general not the double of the summed processes:

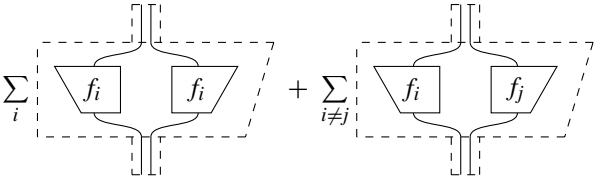


In fact, non-trivial sums of doubled maps are not even **pure quantum maps**, so they can't be obtained by doubling at all.

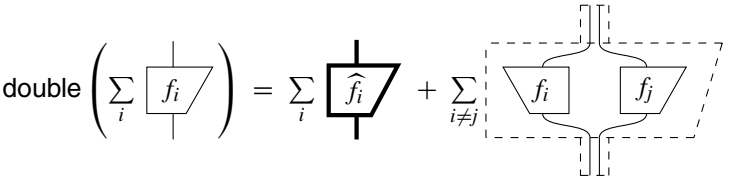
The crucial point is that unfolding the doubling operation yields two independent summations (i.e. over different indices):



We can split this sum into the parts where $i = j$ and where $i \neq j$:



Thus the LHS equals the RHS, plus the 'off-diagonal' terms:



These off-diagonal terms will not go to zero in general. We can already see this for the case of numbers. Letting $\lambda_0 = \lambda_1 = 1$, we have:

$$\mathsf{double}\left(\sum_i \langle\lambda_i\rangle\right) = \mathsf{double}(1+1) = 4 \neq 2 = 1+1 = \sum_i \langle\lambda_i\rangle$$

This is a feature, not a bug. Summing pure quantum maps has an important conceptual meaning, which has no counterpart when summing the underlying linear maps. It can be given a clear physical interpretation as introducing some uncertainty in which process happened. This is called 'mixing' and will be discussed in detail in Section 6.2.7.

The other kind of sums, that is, those that are made in **linear maps** before doubling, gives rise to *quantum superpositions*. We discuss these in Section 7.1.2. So in the end we have two kinds of sums around, each meaning a different thing.

On the other hand, the fact that doubling doesn't preserve sums confirms what we already pointed out in Remark 5.34, namely, that sums are out of place in diagrammatic reasoning, and they may easily cause mistakes. Thus it is prudent to avoid them as much as possible. Over the next two chapters, many concepts will be initially introduced with sums, but gradually replaced by their purely diagrammatic counterparts.

6.1.5.3 Doubled ONB Effects Are 'Quantum Measurements'

If an ONB in **linear maps**:

$$\mathcal{B} = \left\{ \boxed{i} \right\}_i$$

contains at least two states, then:

$$\mathsf{double}\,(\mathcal{B}) := \left\{ \boxed{i} \right\}_i$$

is <u>not</u> a basis in **pure quantum maps**. To see this, consider two states:

$$\boxed{\psi} := \sum_j \boxed{j} \qquad\qquad \boxed{\phi} := \sum_j e^{i\alpha_j} \boxed{j}$$

where the α_j are all distinct. Since ϕ has at least two terms with non-equal coefficients $e^{i\alpha_j}$, ψ and ϕ are not within a global phase of each other, so:

$$\boxed{\widehat{\psi}} \neq \boxed{\widehat{\phi}}$$

However, one can easily verify that for all i:

$$\begin{array}{c} i \\ \boxed{\widehat{\psi}} \end{array} = \begin{array}{c} i \\ \boxed{\widehat{\phi}} \end{array}$$

So $\mathsf{double}(\mathcal{B})$ cannot possibly be a basis.

The reason why $\mathsf{double}(\mathcal{B})$ is not a basis in **pure quantum maps** is that it misses out on all of the 'local' phase information (i.e. the numbers $e^{i\alpha_j}$ above) from ϕ, which (unlike global phases) are highly relevant to the quantum state $\widehat{\phi}$. An important physical consequence of this fact is that *quantum measurements* can extract only a fraction of the information about the state of a system. In the next chapter, we will define quantum measurements and show that measurements defined in terms of ONB effects are in fact the best we can do (though still pretty poor!) when it comes to extracting information from a single quantum state.

At this point, you might start to wonder whether the theory of **pure quantum maps** has bases at all. We saw from Theorem 5.14 that bases allow us to completely characterise processes by a finite set of numbers (namely, its matrix). This is the basis of *quantum tomography*, which as we'll see in the next chapter, is all about identifying a state or a process by means of quantum measurements. The only way this would be possible is if there are still bases around. Thankfully, this is the case.

Theorem 6.24 For any ONB:

$$\mathcal{B} = \left\{ \boxed{j} \right\}_j$$

on a Hilbert space A, the set of states $\mathsf{double}(\mathcal{B})$ can be extended to a (non-orthogonal) basis in **linear maps** for the type $A \otimes A$, consisting entirely of pure quantum states. Hence, in particular this is also a basis in **pure quantum maps** for the doubled system-type $\widehat{A}$.

Proof There are many ways to extend a basis for pure states. For example, let $\mathcal{A}$ be the set of all states of the form:

$$\boxed{\psi_{jk}} := \begin{cases} \frac{1}{2}\left(\boxed{j} + \boxed{k} \right) & \text{if } j \leq k \\ \frac{1}{2}\left(\boxed{j} + i \boxed{k} \right) & \text{if } j > k \end{cases}$$

Each basis state $j \in \mathcal{B}$ is then given as ψ_{jj}, so $\mathcal{B} \subseteq \mathcal{A}$. To show:

$$\mathsf{double}(\mathcal{A}) \;=\; \left\{ \boxed{\psi_{jk}} \; \boxed{\psi_{jk}} \right\}_{jk}$$

is a basis for $A \otimes A$, it suffices to show that any element of a product basis:

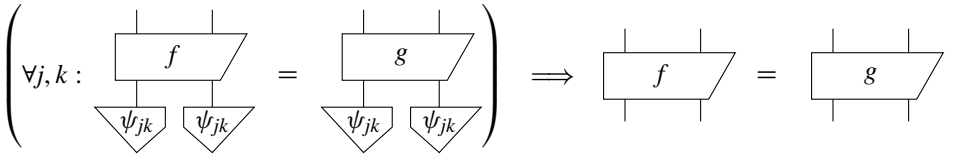
$$\left\{ \boxed{j} \; \boxed{k} \right\}_{jk} \tag{6.12}$$

can be obtained using sums of states of the form $\lambda\psi_{jk}$ for $\lambda \in \mathbb{C}$, which is left as an exercise to the reader. □

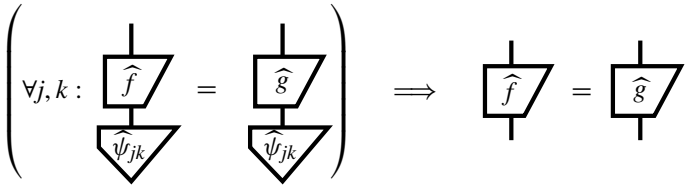
So, we can still distinguish processes by how they act on states, but it's not nearly as easy as it used to be! In particular, there is no way to build an ONB of pure quantum states.

Convention 6.25 Even though Theorem 6.24 implies that the 'dimension' (in the sense of Definition 5.2) of a quantum system $\widehat{A}$ is D^2 when the dimension of A is D, when we talk about the dimension of a quantum system, we are always referring to the dimension of the (undoubled) Hilbert space. For instance, we still refer to $\widehat{\mathbb{C}^2}$ as a two-dimensional quantum system.

Remark 6.26 In the statement of the previous theorem we distinguished between $\mathsf{double}(\mathcal{A})$ being a basis in **linear maps** and being a basis in **pure quantum maps**. These are indeed two different things. A basis in **linear maps** needs to uniquely fix all linear maps via:



whereas a basis for **pure quantum maps** only needs to uniquely fix linear maps of the form $\widehat{f}, \widehat{g}, \ldots$ via:



which is a weaker requirement. We will use the full strength of Theorem 6.24 in Section 6.2.1.2.

6.2 Quantum Maps from Discarding

In the previous section we transformed **linear maps** into **pure quantum maps** via doubling. So far, doubling was merely a way of recasting an existing thing into something derived, which we depicted by showing it in boldface. However, the new types are made from two wires rather than one, and since two is more than one, there is actually some extra space for genuinely new stuff, which has no counterpart in the underlying process theory of **linear maps**. We will now show that we are missing out on a very important process, namely the process of *discarding* a system, which is hiding in this extra space. In fact, this is the only extra thing we need to throw in to get all **quantum maps**.

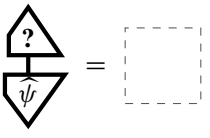
6.2.1 Discarding

When we introduced effects back in Section 3.4.1, the first example we gave was 'discarding a system'. Here, 'discarding a system' may mean ignoring it, destroying it, or maybe firing it off into space, never to be seen again. Alternatively, it could be the case that there is some piece of a bigger system that we simply do not have access to, so we can think of it as being 'discarded' for all intents and purposes.

In this section, we will home in on discarding, by looking at its behaviour and showing that this behaviour forces us to make one particular choice. We begin with the following realisation.

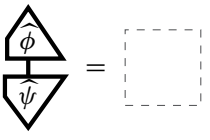
6.2.1.1 Discarding Is Not a Pure Quantum Map

To see this, it suffices to consider how a discarding process should behave on the state of a single system. Discarding should do nothing but remove that state from the picture:



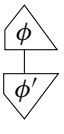
or put another way: it is a test that succeeds with certainty, but otherwise tells us nothing about the state. In particular, the discarding process cannot depend on the state of the system that gets discarded.

Proposition 6.27 For non-trivial Hilbert spaces, i.e. with dimension > 1, there exists no pure quantum effect $\widehat{\phi}$ such that for all normalised pure quantum states $\widehat{\psi}$ we have:

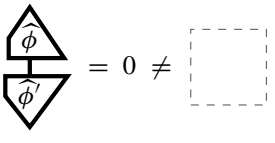


(6.13)

Proof Suppose that $\widehat{\phi}$ is a pure discarding effect. Clearly $\widehat{\phi}$ cannot be zero. Thus there exists some λ such that $\lambda\phi$ is normalised. By Theorem 5.79, $\lambda\phi$ extends in an ONB. Since the dimension is at least two, let ϕ' be a distinct state in that ONB. It must therefore be orthogonal to $\lambda\phi$, and hence also orthogonal to ϕ. Consequently:

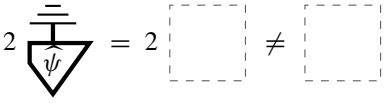


and hence:



which is a contradiction. □

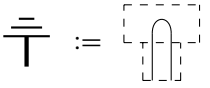
Remark 6.28 In Section 4.3.3, we gave a good reason for considering the normalised states by default, namely that these are precisely the states that return 'yes' with certainty when tested for themselves. To even get a discarding map in the first place, this restriction is essential. If for some $\widehat{\psi}$ we have (6.13), then for $2\widehat{\psi}$ this would no longer be the case:



so (6.13) would simply be impossible to satisfy.

Since we cannot use any pure effect, we propose the following instead.

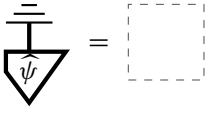
Definition 6.29 We define *discarding* to be the effect:



(6.14)

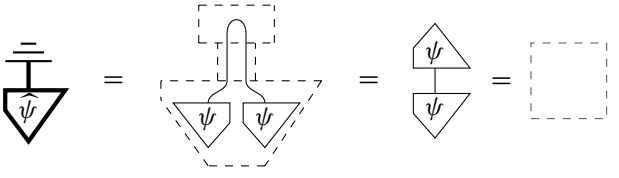
This effect indeed behaves as required.

Proposition 6.30 For any normalised pure quantum state $\widehat{\psi}$ we have:



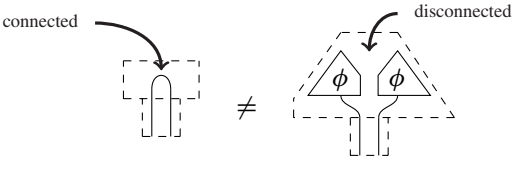
(6.15)

Proof Since $\widehat{\psi}$ is normalised, so is ψ (cf. Exercise 6.22), so we have:



□

Discarding as defined in (6.14) certainly doesn't look like a pure quantum effect:



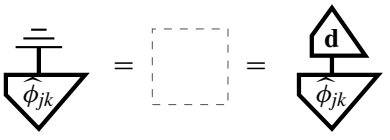
which is consistent with Proposition 6.27. It also behaves as expected on states, so it looks like we're done. But before we hang up our hats, it's worth asking whether this is our only choice for a discarding effect (spoiler alert: yes!).

6.2.1.2 There Is Only One Linear Map Fit for Purpose

Theorem 6.31 The discarding map defined in Definition 6.29 is the unique linear map sending all normalised pure quantum states to 1.

Proof Suppose there exists some other effect:

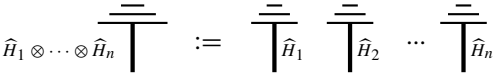
that sends all normalised pure states to 1. From Theorem 6.24, we saw that there exists a basis $\mathsf{double}(\mathcal{B}')$ of pure quantum states for $A \otimes A$ in **linear maps**. Let $\mathsf{double}(\mathcal{B}'')$ be the basis formed by normalising each of the states in $\mathsf{double}(\mathcal{B}')$ (which is clearly still a basis). Then, for all $\widehat{\phi}_{jk} \in \mathsf{double}(\mathcal{B}'')$:



Since **d** and discarding agree on a basis, they must therefore be equal. Hence discarding is unique. □

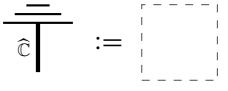
Now that we know that discarding is uniquely defined by its intended behaviour, we can derive what it should be in certain special cases.

Exercise 6.32 Show that:



(6.16)

and that:



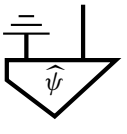
(6.17)

(noting that $\widehat{\mathbb{C}}$ is the 'no wire' system for **pure quantum maps**).

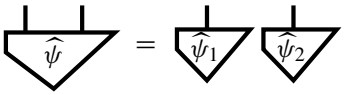
6.2.1.3 Discarding Does Not Preserve Pure Quantum States

If we start with a pure quantum state on two systems and discard one system, the resulting state typically won't be a pure quantum state. In fact, the only case where it will be a pure state is when we start with something that is $\otimes$-separable.

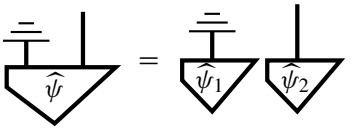
Proposition 6.33 For any pure quantum state $\widehat{\psi}$, the *reduced state*:



is a pure state if and only if $\widehat{\psi}$ is $\otimes$-separable:

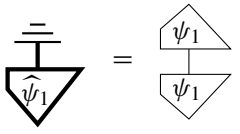


Proof For ($\Leftarrow$), we have:

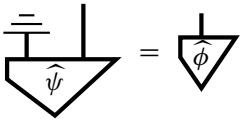


(6.18)

In the proof of Proposition 6.30 we already saw that:

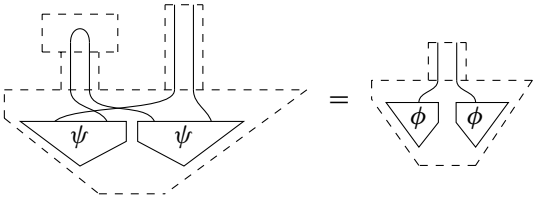


which is a positive number. Since every positive number is pure (cf. Proposition 6.23), the state (6.18) is a pure quantum state. For ($\Rightarrow$), assume there exists some pure state $\widehat{\phi}$ such that:

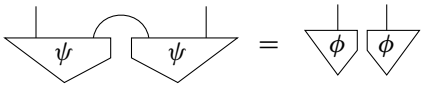


(6.19)

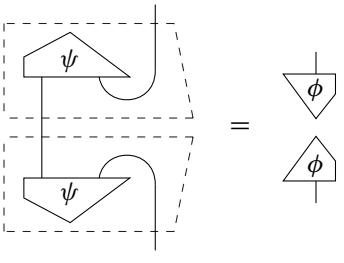
We will rely on Proposition 5.74, which states that $f^\dagger \circ f$ is $\circ$-separable if and only if f is. Unfolding the doubled maps in (6.19):



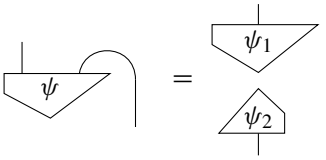
we obtain:



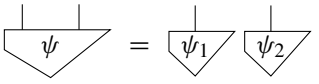
By process–state duality and using transposition we get:



Then, by Proposition 5.74, there exist ψ_1 and ψ_2 such that:



(Note that there is no loss of generality by depicting the effect ψ_2 in conjugate form.) Thus, again using process–state duality, ψ is $\otimes$-separable:

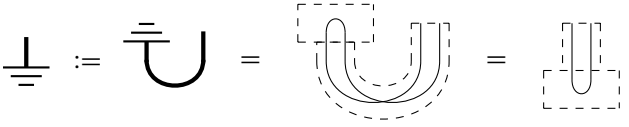


Doubling the equation above yields the required condition. □

As reduced states are not, in general, pure quantum states, we need to introduce a more general family of states to account for the fact that parts of systems may be discarded.

6.2.2 Impurity

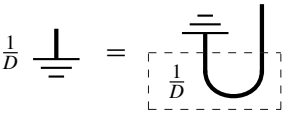
When we compose arbitrary pure quantum maps with the discarding map, lots of new stuff emerges. For example, consider the transpose (or equivalently, the adjoint) of the discarding map:



This state is so important that its normalised version gets a special name.

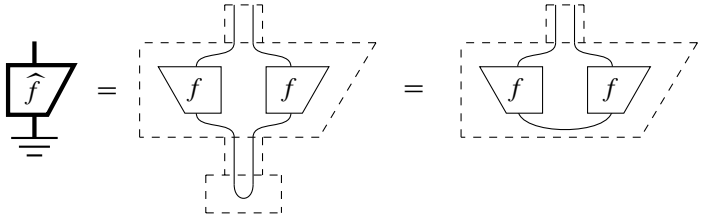
Definition 6.34 The *maximally mixed state* is:

This maximally mixed state is an example of a reduced state (cf. Proposition 6.33), namely what's left after discarding half of a Bell state:

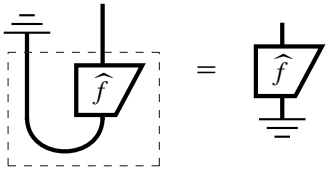


As we mentioned in the introduction, the term 'mixed' has to do with a lack of knowledge about the actual state of the system. In Section 6.2.7 we'll see that the maximally mixed state means we have no knowledge whatsoever about the state of a system.

We can now generate other new quantum states either by applying a pure quantum map to a system in a maximally mixed state:

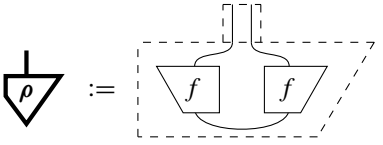


or, equivalently, by discarding one system of a bipartite pure state (represented here by means of process–state duality):



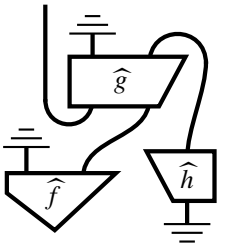
The resulting form is moreover totally generic.

Theorem 6.35 Any state obtained by composing pure quantum maps and discarding is a *quantum state*, i.e. a state of the form:

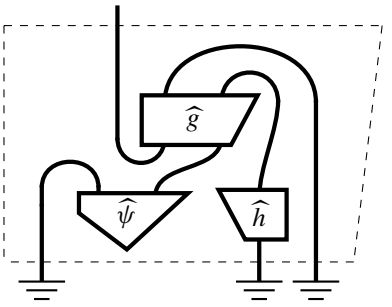


(6.20)

Proof For a diagram consisting of some pure quantum maps and discarding:



we can always pull all of the discarding maps (or maximally mixed states) down to the bottom, using caps as necessary:

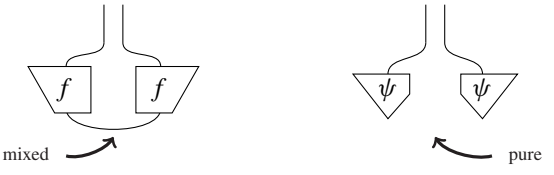


We can then combine all of the maximally mixed states into a single state (cf. equation (6.16) upside-down). Thus, we obtain a pure quantum map, applied to a maximally mixed state. □

Recalling the definition of ⊗-positivity from Section 4.3.6, we obtain the following.

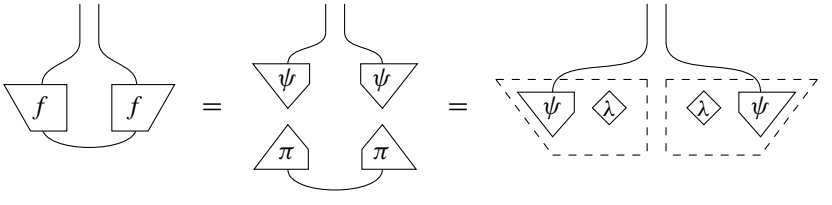
Corollary 6.36 Quantum states are ⊗-positive states in **linear maps**.

The form of quantum states generalises that of pure quantum states, where the map f in (6.20) has a trivial input wire. In particular, not all quantum states are pure states. When they fail to be pure states, we call them *impure*. Unfolding the form of an impure quantum state $\boldsymbol{\rho}$ and a pure one $\widehat{\psi}$ the difference comes down to the presence of a wire connecting the left half to the right half:



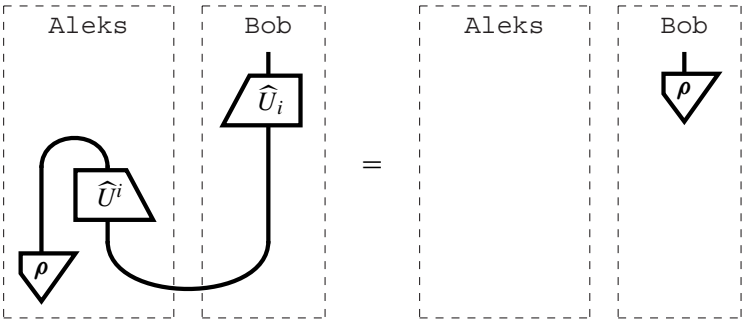
So *purity* itself is a diagrammatic notion.

Remark 6.37 Note that we should take the absence of a wire as evidence for purity. Conversely, the presence of a wire only indicates the possibility of being impure. For example, suppose f itself is disconnected:

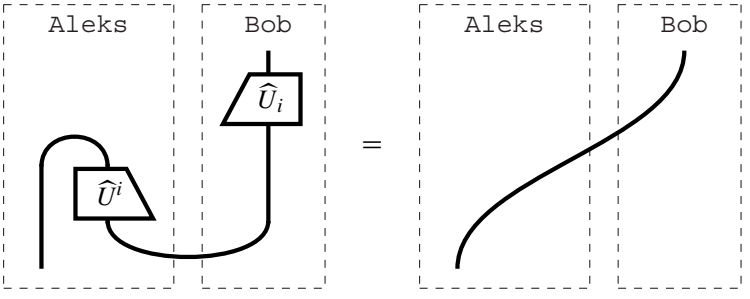


with $\lambda := \sqrt{\pi^\dagger \circ \pi}$. Then the resulting state is pure.

Example 6.38 In quantum teleportation it doesn't matter if the state that we teleport is impure, since the picture stays exactly the same:



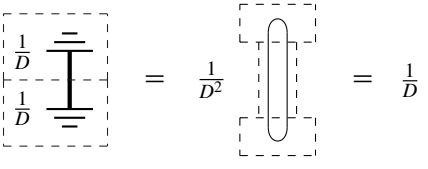
While this is a proper generalisation, the reason that it doesn't require any additional work is the compositional nature of string diagrams. The key to the above equality being true is the fact that have:



Then, clearly it doesn't matter whether we plug in $\widehat{\psi}$ or ρ at the input.

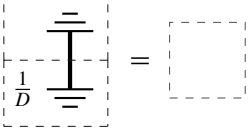
6.2.3 Weight and Causality for Quantum States

Now, the cautious reader may have noticed something sneaky about Definition 6.34, namely, that the maximally mixed state is not normalised in the sense of Definition 4.48:



(6.21)

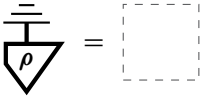
since the circle is equal to the dimension D, as we saw in Corollary 5.33. On the other hand, if we discard the maximally mixed state we do get:



So unlike in the case of pure quantum states, where in Proposition 6.30 we showed that the squared norm and the number arising from discarding coincide, this ceases to be the case for the maximally mixed state and, more generally, will no longer hold for impure quantum states. This justifies the introduction of a new name for the numbers obtained when discarding.

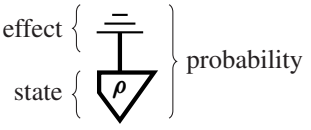
Definition 6.39 The *weight* of a quantum state ρ is:

and ρ is *causal* if its weight is 1, i.e.:

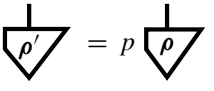


(6.22)

How should we interpret this quantity? The Born rule tells us that:



So, the weight is the result of performing a trivial test on the state (i.e. 'is this a state?'). Normally, we would expect this test to return 'yes' with probability 1. However, we will see later in this chapter that a state may be the result of a non-deterministic process. In this case, the weight then tells us what the probability is to end up in this state. Causal states are then those that occur with certainty. In other words, what we call a quantum state is really a combination of two things:



an 'actual' state of a system (the causal state ρ) and some probability p that it occurred (i.e. the weight of ρ'). We continue this discussion in Section 6.4.1 when we introduce non-deterministic quantum processes properly.

Ignoring non-determinism for the moment, the only 'actual' states are the causal ones. So, we can more fundamentally interpret the causality equation as follows:

If a state is discarded, it may as well never have existed.

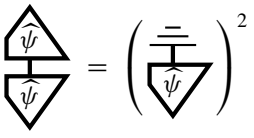
This is obviously a reasonable, and furthermore necessary, assumption to make. There are many systems out there (e.g. on Mars) that we have no control over and know nothing about. So, we ignore (i.e. discard) them in our calculations. If we weren't allowed to do so, we pretty much couldn't do any science. There is also a hint of relativity theory here: if something is sufficiently far away (i.e. or 'space-like separated', to use relativity lingo),

it should not be able to affect what is happening locally, because light simply cannot travel fast enough. This feature, called 'non-signalling', is fundamental to relativity theory. In Section 6.3, we will see how non-signalling and causality are closely connected.

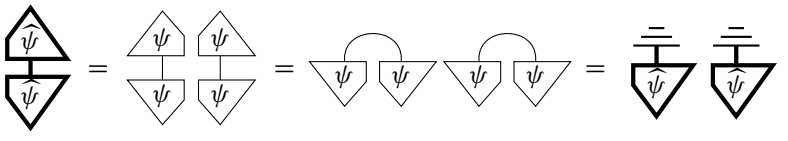
For states, the constraint imposed by causality is fairly weak: it restricts the weight to 1. Hence, any non-zero state can be turned into a causal one simply by composing it with the appropriate number. However, we shall see in Section 6.2.6 that for arbitrary processes the constraint is more drastic, and, as we will show in Theorem 6.54, for effects it is pretty extreme!

For pure states, we already know that normalisation and causality coincide. More generally, we have the following.

Proposition 6.40 For any pure state $\widehat{\psi}$:



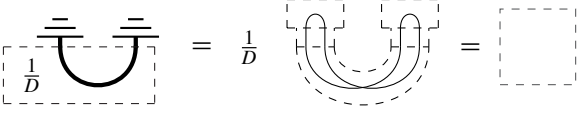
Proof This can be shown by unfolding, then transposing the effects:



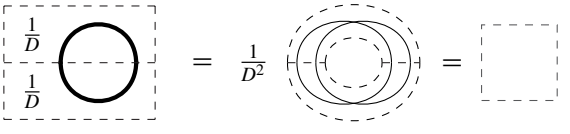
□

The fact that normalisation and causality coincide for pure states is then just the special case where the squared norm and the (squared) weight are both equal to 1.

Example 6.41 Since the circle is equal to D, the Bell state is causal:

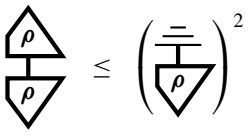


Given that the Bell state is moreover pure, it is also normalised:



However, as we already saw in (6.21), the squared norm can be smaller than the squared weight for impure states. The fact that these two quantities do not always coincide is actually useful, since it gives us an effective way to figure out when a state is pure.

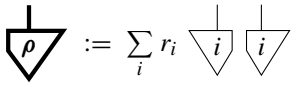
Proposition 6.42 For any quantum state ρ we have:



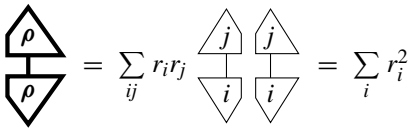
(6.23)

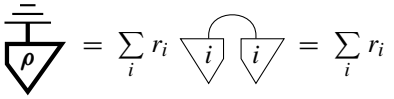
and we have equality if and only if ρ is pure.

Proof Since ρ is a $\otimes$-positive state, by the spectral theorem (and Corollary 5.72 in particular) there exists some ONB and positive numbers r_i such that:

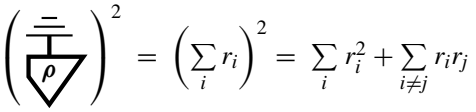


It is then straightforward to compute the squared norm and the weight:





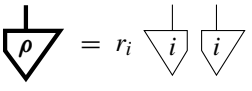
Thus:



Since all $r_i r_j \geq 0$ the first claim follows. If ρ is pure, then by Proposition 6.21, the squared norm and the squared weight coincide. Conversely, assume the squared norm and the squared weight are equal. Then:

$$\sum_{i \neq j} r_i r_j = 0$$

which is only true if, for all $i \neq j$, we have $r_i r_j = 0$. In that case, at most one r_i is non-zero, so:

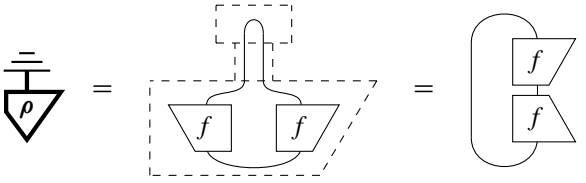


which is a pure quantum state. □

Proposition 6.42 does much more than provide a means of detecting whether a state is pure. As a causal state becomes more and more impure, the squared norm will go lower and

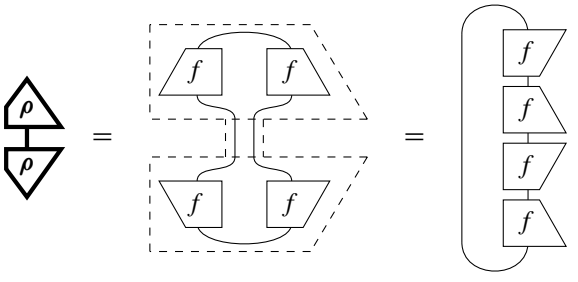
lower, until it hits $1/D$ for the maximally mixed state. Thus, it is a quantity that actually measures the impurity of a state. We will explore this and other measures of impurity in Section 13.2.

We can also visualise the difference between the quantities in (6.23). For the weight we have:



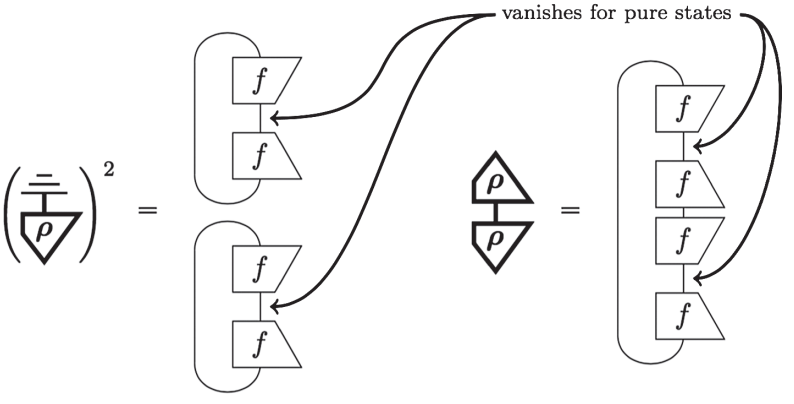
(6.24)

and for the squared norm we have:



(6.25)

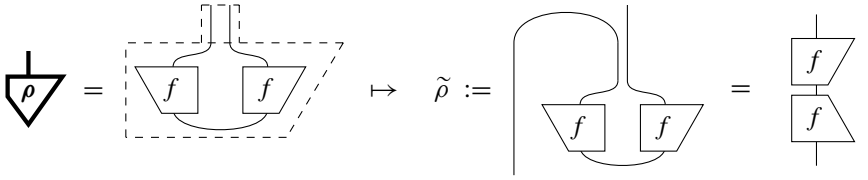
In the case of a pure state it is the absence of the following wires that results in the squared weight and the squared norm becoming the same:



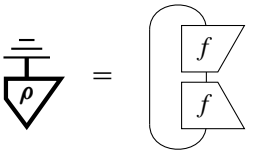
Otherwise, it is the difference in wiring that causes the difference in numbers.

Remark 6.43 Recall from Remark 6.5 that we could translate between a pure state and its representation as a density operator, which in that case was a projector. This generalises to (causal) mixed states, whose associated density operator is a positive map that has trace 1.

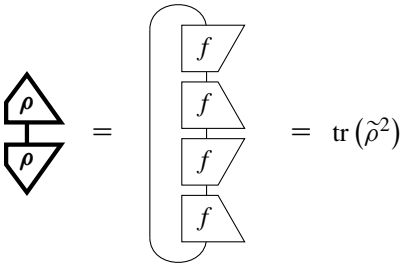
The translation between our doubled representation and the density operator representation is provided by process–state duality:



In the density operator representation, from (6.24) it follows that discarding a state means taking its trace:



so a density operator with trace 1 corresponds to a causal state. Similarly, from (6.25) it follows that:



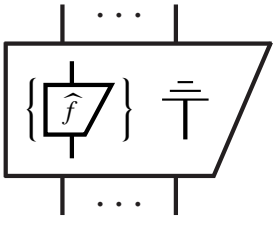
and so (6.23) becomes, in terms of density operators:

$$\mathrm{tr}\left(\widetilde{\rho}^2\right) \leq (\mathrm{tr}\,(\widetilde{\rho}))^2$$

6.2.4 The Process Theory of Quantum Maps

We passed from pure quantum states to quantum states by adjoining discarding. In fact, the entire process theory of quantum maps is also obtained in that manner.

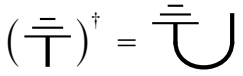
Definition 6.44 The process theory of **quantum maps** has as types doubled Hilbert spaces $\widehat{A}$ and as processes all diagrams made from pure quantum maps and discarding:



and the following should come as no surprise.

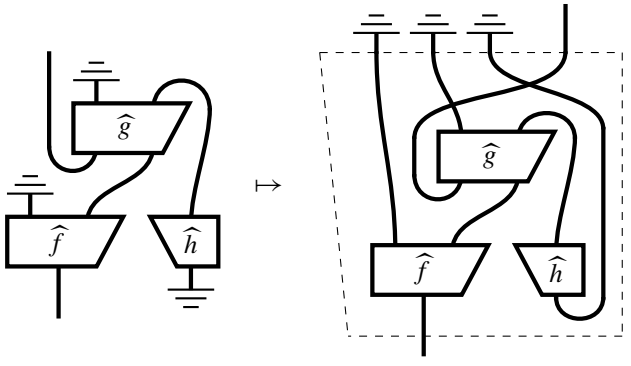
Theorem 6.45 The theory of **quantum maps** admits string diagrams.

Proof **Quantum maps** inherit their caps and cups from **pure quantum maps**, so it suffices to show that **quantum maps** have adjoints. The adjoint of a pure quantum map is another pure quantum map, and the adjoint of discarding is just its transpose:



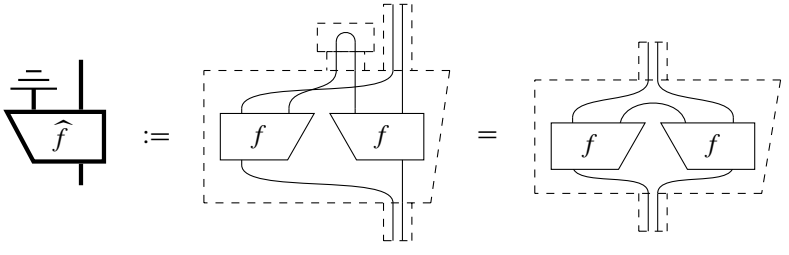
which is the composition of a pure quantum map (the cup) with discarding. Thus it is again a quantum map. Since adjoints need to preserve diagrams and all diagrams in **quantum maps** are made up of pure quantum map and discarding, every quantum map has an adjoint. □

Just as we saw with quantum states in the previous section, we can put any quantum map in a 'normal form' by grouping all of the discarding maps together into a single effect:



Hence we have the following.

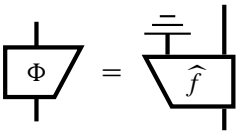
Proposition 6.46 All quantum maps are of the form:



for some linear map f.

So, quantum maps are precisely those linear maps that are $\otimes$-positive processes (cf. Definition 4.66). Proposition 6.46 also shows that arbitrary quantum maps arise by ignoring

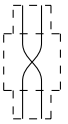
part of the output of a process. In other words, for any quantum map Φ, we know that there must exist a pure quantum map $\widehat{f}$ such that:



(6.26)

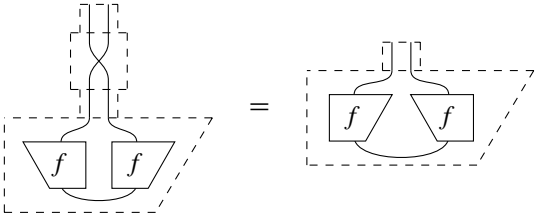
Definition 6.47 We refer to the pure quantum map $\widehat{f}$ in (6.26) as a *purification* of the quantum map Φ.

It might be tempting to think that quantum maps are precisely those maps that send quantum states to quantum states, but this is not the case. For example, the swap linear map, when conceived as a linear map from type $\widehat{A}$ to $\widehat{A}$:



(6.27)

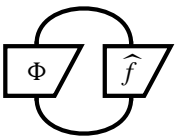
clearly sends quantum states to quantum states:



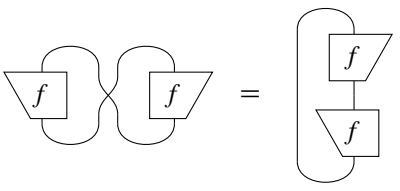
(it actually conjugates them!) However ...

Proposition 6.48 The linear map (6.27) is not a quantum map.

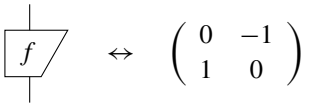
Proof For any quantum map Φ and any linear map f the following number:



will also be a quantum map, namely a positive number. However, if instead of Φ we take (6.27), we obtain:



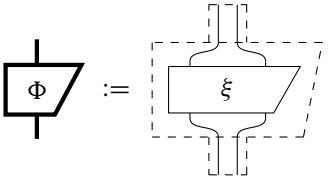
This number is not of the form of a linear map composed with its adjoint, so it cannot be positive for all f. And indeed, picking:



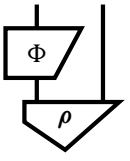
yields -2, which is a contradiction. □

Being a quantum map is actually stronger than just preserving states of a single system. The reason is that quantum maps should not only be well behaved in that special case but should be well behaved when they are included in any diagram. For example, applying a quantum map to just one part of a state on multiple systems, as in (6.28) below, should again yield a quantum state. In fact, this (seemingly more specific) condition is actually equivalent to being a quantum map.

Theorem 6.49 A linear map:



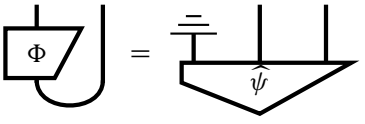
is a quantum map if and only if for all quantum states ρ we have that:



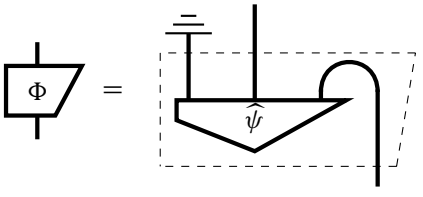
(6.28)

is again a quantum state.

Proof If Φ is a quantum map, (6.28) must be a quantum state simply because **quantum maps** is a process theory. For the other direction, let ρ in (6.28) be the doubled cup state. Then by assumption, this is a quantum state, so by Proposition 6.46, there exists a pure quantum state $\widehat{\psi}$ such that:

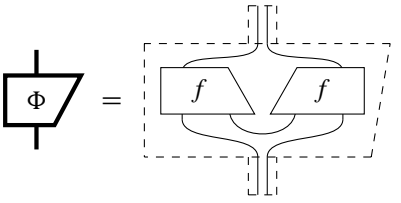


So by process–state duality it follows that:



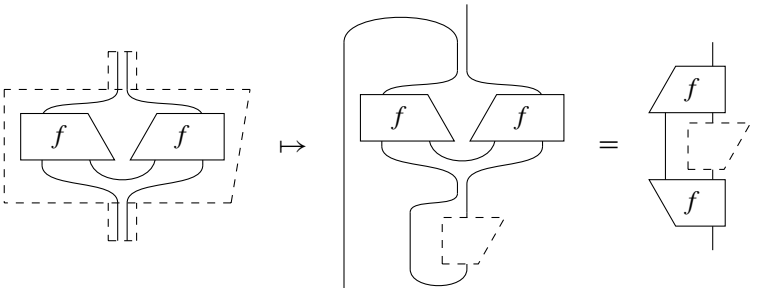
and, hence, Φ is indeed a quantum map. □

Remark 6.50 We pick things up where we left off in Remark 6.43. First note that we can equivalently represent a general quantum map as follows:

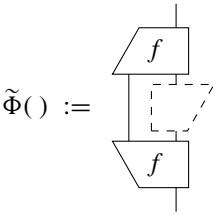


(6.29)

We can then reshape it a bit:

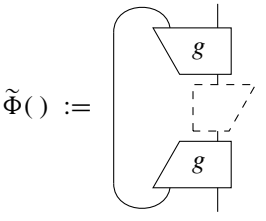


Here, the small dashed box represents a hole where we can insert a density operator $\widetilde{\rho}$ and obtain $\widetilde{\Phi}(\widetilde{\rho})$. The 'super-operator' (i.e. a map taking operators to operators):



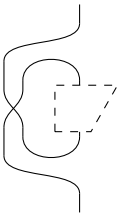
(6.30)

is usually called a *completely positive map* (CP-map). They are also commonly seen in an equivalent form, involving the trace:



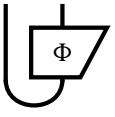
The 'complete' part refers to the fact that a CP-map preserves positivity (i.e. 'being a quantum state') even when applied to just part of a system, as in (6.28). The super-operator

analogue to preserving positivity only when applied to entire systems is called simply *positive*. In the case of (6.27) the positive super-operator looks like this:



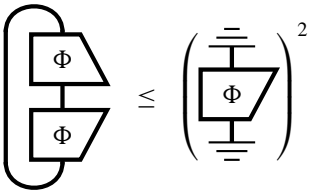
which, by Proposition 6.48, is not of the form (6.30).

When specialising Proposition 6.42 to the case of a bipartite state:



we now obtain a means of determining if a quantum map is pure, as well as a quantity that represents the degree of impurity for general quantum maps.

Corollary 6.51 For any quantum map Φ we have:



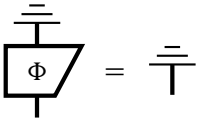
(6.31)

and we have equality if and only if Φ is pure.

6.2.5 Causality for Quantum Maps

We now generalise the definition of causality from states to maps.

Definition 6.52 We call a quantum map Φ *causal* if we have:

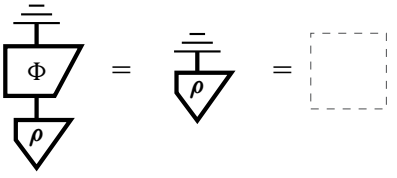


(6.32)

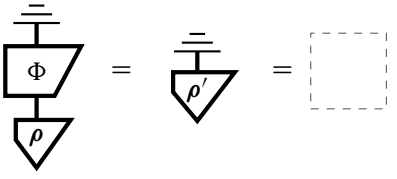
In fact, causality for quantum maps means preservation of causal states.

Proposition 6.53 A quantum map is causal if and only if it sends causal quantum states to causal quantum states.

Proof For a causal quantum map Φ and a causal quantum state ρ we have:

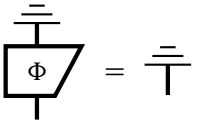


so ($\Rightarrow$) indeed holds. For ($\Leftarrow$), suppose Φ sends any causal state ρ to another causal state ρ'. Then we have:



Thus the following effect:

sends all causal states (and in particular, all normalised pure states) to 1. Then, by Theorem 6.31 on uniqueness of discarding:



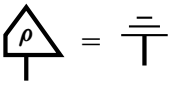
so the quantum map Φ is indeed causal. □

We can also interpret (6.32) directly:

If the output of a process is discarded, it may as well have never happened.

This is a straight generalisation of the interpretation we gave for causal states in Section 6.2.3.

Despite seeming innocent at first, causality has a somewhat shocking consequence when applied to effects. By equation (6.17), discarding the 'no wire' system is the same as doing nothing. Since effects have no outputs, causality reduces to this equation:



which forces any causal effect to be equal to discarding!

Theorem 6.54 There is a unique causal quantum effect: discarding.

So causal effects in **quantum maps** are utterly uninteresting. If causality is going to play such an important role, why did we even bother to introduce effects at all? Don't worry, your time hasn't been wasted! Once we consider non-deterministic quantum processes, we will be able to realise all quantum effects non-deterministically, and this will be vital for applications such as quantum teleportation.

Remark 6.55 The analogue to causality for CP-maps (cf. Remark 6.50) is that they are trace preserving:

$$\mathrm{tr}\,(\widetilde{\Phi}(\widetilde{\rho})) \;=\; \mathrm{tr}\,(\widetilde{\rho})$$

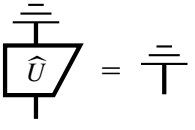
since trace-preserving CP-maps send density operators to density operators, just as causal quantum maps send causal quantum states to causal quantum states, as we saw in Proposition 6.53.

6.2.6 Isometry and Unitarity from Causality

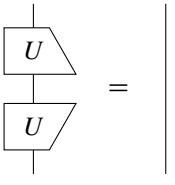
So what is the constraint imposed by causality for pure quantum maps? We already saw in Section 6.2.3 that for pure states causality means normalisation. Realising that a normalised state is just a special case of an isometry, when the input system is trivial, we can generalise this statement.

Theorem 6.56 For pure quantum maps, the following are equivalent:

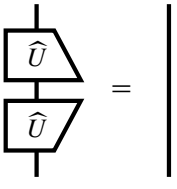
1. $\widehat{U}$ is causal:



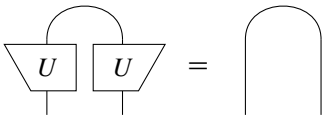
2. U is an isometry:



3. $\widehat{U}$ is an isometry:

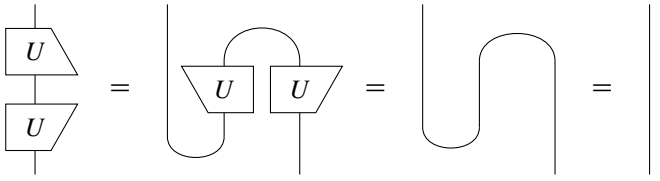


Proof Unfolding the causality equation, we have:



(6.33)

so we obtain the isometry equation by bending the left input wire up:



and hence we obtain $1 \Leftrightarrow 2$. By Theorem 6.20 we obtain $2 \Leftrightarrow 3$. □

We already saw above in Theorem 6.54 that discarding, which of course is not pure, is the only causal quantum effect. Hence, there are no pure causal quantum effects. More generally, Theorem 6.56 implies that there exist no pure causal quantum maps from $\widehat{A}$ to $\widehat{B}$ if $\dim(A) > \dim(B)$, because in that case there exist no isometries from A to B (cf. Proposition 5.81).

Now recall from Proposition 5.81 that any isometry from a Hilbert space to itself must be a unitary. This yields an easy corollary to Theorem 6.56.

Corollary 6.57 A pure quantum map from a system $\widehat{A}$ to itself is causal if and only if it is a unitary.

Of course, this fact depends on the dimension theorem, which we have only established for **linear maps** in particular. On the other hand, Theorem 6.56 does not rely on any special properties of linear maps. A similarly general consequence is the following. By Proposition 4.58, invertible isometries are unitary. Combining this with Theorem 6.56 thus yields the following.

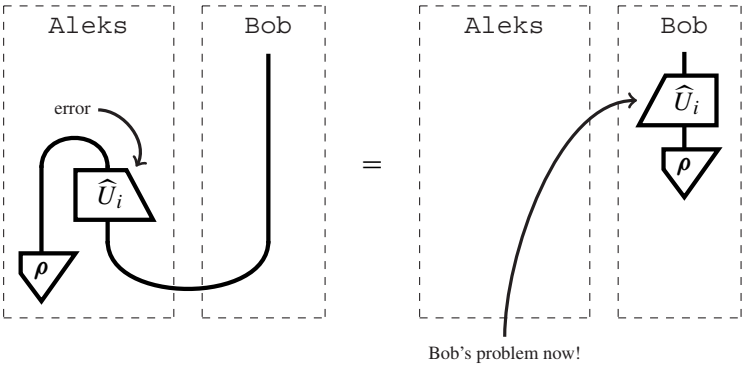
Corollary 6.58 For a pure quantum map, the following are equivalent:

1. It is causal and invertible.
2. It is unitary.

Remark 6.59 In many textbooks, unitarity is assumed from the start, without (much) justification. However, causality, with its simple physical interpretation, is much easier to justify. Thus, a pleasant consequence of doubling is that isometry (and hence unitarity) fall out so easily, as in equation (6.33).

We already made implicit use of the unitarity requirement.

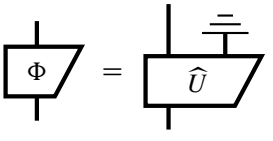
Example 6.60 Recall that in Section 4.4.4 where we presented quantum teleportation we assumed that $\widehat{U}_i$ needed to be unitary in the RHS of:



in order for Bob to be able to fix the error. Now we know why: in order for Bob's correction to be causal, it needs to be an isometry, and in order to undo $\widehat{U}_i$ it must furthermore be unitary.

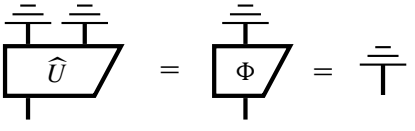
So what about impure processes? If we combine the fact that all causal pure quantum maps are isometries with the fact that all quantum maps can be purified, we can immediately see that every causal quantum map can be represented as an isometry with one of its outputs discarded.

Theorem 6.61 (Stinespring dilation I) For every causal quantum map Φ there exists an isometry $\widehat{U}$ such that:



(6.34)

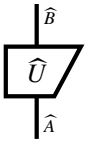
Proof By Proposition 6.46 we know that there always exists a pure quantum map $\widehat{U}$ such that 6.34 holds. By causality of Φ, it follows that $\widehat{U}$ must also be causal:



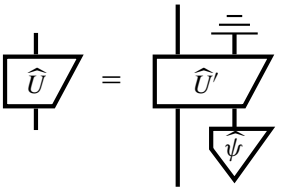
which, by Theorem 6.56 implies that $\widehat{U}$ is an isometry. □

In fact, we can boost this result from isometries to unitaries, but this requires a bit more work. First, note that we can replace any isometry with a unitary and a pure state.

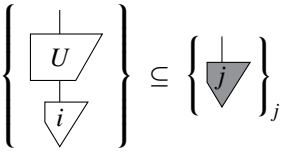
Lemma 6.62 For any isometry U:



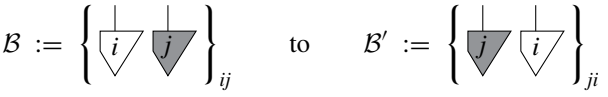
there exists a unitary U' and a pure quantum state $\widehat{\psi}$ such that:



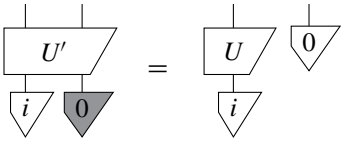
Proof Fix ONBs for A and B such that:



Then, fix a linear map U' from $A \otimes B$ to $B \otimes A$, which sends:

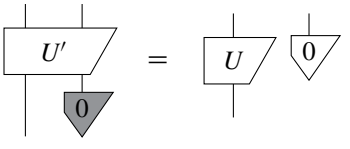


such that:

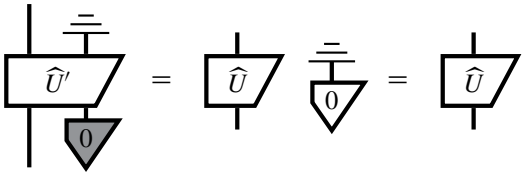


(6.35)

and the remaining basis states of $\mathcal{B}$ are mapped (injectively) to the remaining basis states of $\mathcal{B}'$. This is always possible since $\dim(A \otimes B) = \dim(B \otimes A)$, and yields a bijection on the ONBs $\mathcal{B}$ and $\mathcal{B}'$. Thus by Proposition 5.38, U' is unitary. Since equation (6.35) holds for all white ONB-states, we have:



Doubling this equation and discarding the first output yields:

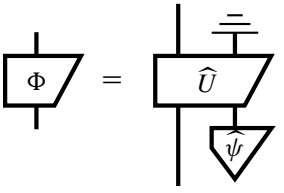


which completes the proof, for:

□

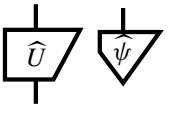
Now we can conclude the following.

Corollary 6.63 (Stinespring dilation II) Every causal quantum map Φ arises from some unitary $\widehat{U}$ by plugging some pure causal quantum state $\widehat{\psi}$ into one of its inputs and discarding one of its outputs:



(6.36)

Remark 6.64 Stinespring dilation is used by some people to justify a point of view where the only real quantum processes are just the pure, unitary ones, of which we only have access to some small part. This belief is sometimes referred to as the 'Church of the larger Hilbert space'. While this point of view is perfectly consistent with quantum theory, it is not very convenient for thinking process theoretically. For example, one cannot build a process theory that includes both unitaries and pure quantum states without also considering general isometries. Indeed, unitary quantum maps U and pure quantum states ψ yield (non-unitary) isometries when composed in parallel:

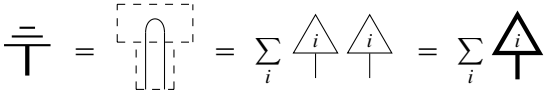


We will return to this point in Section 7.3.2 when we discuss von Neumann's formulation of quantum theory.

6.2.7 Kraus Decomposition and Mixing

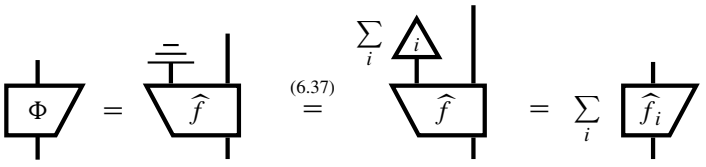
One way to understand impure quantum maps, by their very definition, is in terms of discarding parts of a larger system. We now give an alternative interpretation in terms of *mixing*. For now, this will involve explicit sums. However, in Section 8.3.4 mixing will be given a purely diagrammatic treatment, as part of our general strategy for eliminating sums.

A first step towards mixing is to replace the discarding map by a sum over an ONB. We can do this by decomposing the cap with equation (5.37):

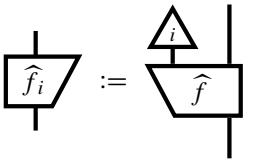


(6.37)

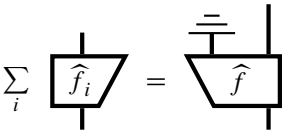
Combining this with purification, we can write any quantum map as a sum of pure quantum maps:



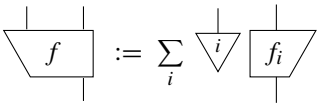
where:



Conversely, the sum of any finite set of pure quantum maps is a quantum map (recall that sums only exist for processes of the same type):



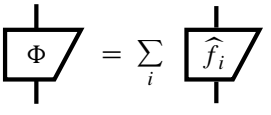
where:



(6.38)

So, in summary, we have the following.

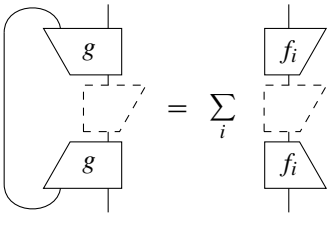
Theorem 6.65 The sum of pure quantum maps is a quantum map, and any quantum map Φ can be written as a sum of pure quantum maps:



(6.39)

Such a representation is known as a *Kraus decomposition*.

Remark 6.66 For completely positive maps (cf. Remark 6.50), a Kraus decomposition becomes:

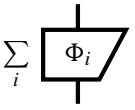


which one usually encounters in the following form:

$$\widetilde{\Phi}(\widetilde{\rho}) := \sum_i f_i \, \widetilde{\rho} \, f_i^\dagger$$

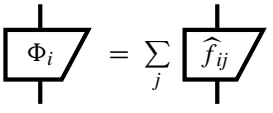
The sum of pure quantum maps is almost never pure, so the theory of **pure quantum maps** isn't closed under sums. However, we have the following theorem.

Theorem 6.67 The sum of any finite set of quantum maps:

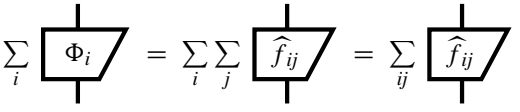


is a quantum map; i.e. the theory of **quantum maps** is *closed under sums*.

Proof Since all of the Φ_i have Kraus decompositions:



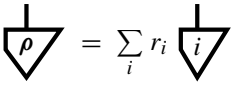
then we can expand their sum as:



which by Theorem 6.65 is a quantum map. □

In Theorem 6.65 we said 'a' Kraus decomposition since Kraus decompositions are not unique. There are, however, some special ones. For Kraus decompositions for quantum states, by the spectral theorem (and Corollary 5.72 in particular), we have the following.

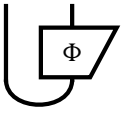
Corollary 6.68 Every quantum state ρ has a Kraus decomposition of the following form:



(6.40)

for some ONB and positive real numbers r_i.

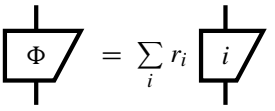
If we apply Corollary 6.68 to:



(6.41)

we can decompose (6.41) over an ONB of biparite states. Equivalently, we can decompose Φ itself over a Hilbert–Schmidt ONB (cf. Definition 5.101).

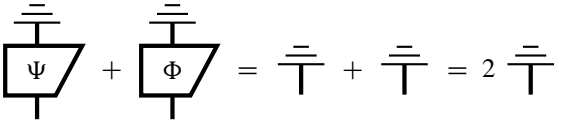
Corollary 6.69 Every quantum map Φ has a Kraus decomposition of the following form:



(6.42)

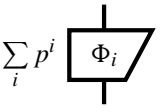
for some Hilbert–Schmidt ONB and positive real numbers r_i.

We showed in Theorem 6.67 that any sum of quantum maps is again a quantum map. Of course, if we add together causal quantum maps, what we get will no longer be causal:



However, instead of ordinary sums, we can consider *convex combinations* of causal quantum maps.

Definition 6.70 A *convex combination* or *mixture* of a family of causal quantum maps $\{\Phi_i\}_i$ is a sum of the form:



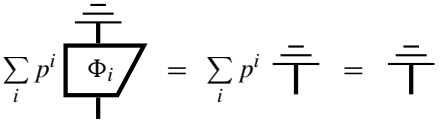
(6.43)

where the numbers p^i sum to 1.

In this case causality is preserved.

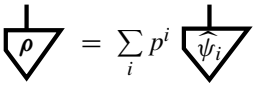
Theorem 6.71 Every convex combination of causal quantum maps is again a causal quantum map.

Proof By Theorem 6.67, the map (6.43) is a quantum map, and by causality of each of the quantum maps Φ_i we have:



□

We refer to the operation that produces (6.43) out of $\{p^i\}_i$ and $\{\Phi_i\}_i$ as *mixing*. For a mixture we can interpret each p^i as the probability that the process Φ_i happens. In other words, there is a lack of knowledge – represented by the probability distribution $\{p^i\}_i$ – about which pure process Φ_i out of the $\{\Phi_i\}_i$ is happening. For example, the causal quantum state:



(6.44)

can be interpreted as a system that is in one of the pure states $\widehat{\psi}_i$, but we don't know which one. We only know the probability p^i that it is in the i-th state. We can in fact write any causal quantum state in the following form.

Theorem 6.72 Every causal quantum state can be regarded as a mixture of pure causal quantum states. Moreover, these pure causal quantum states can always be chosen to form an ONB.

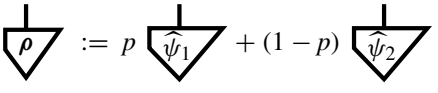
Proof Since ONB states are always causal, this follows immediately from Corollary 6.68. □

So, it is tempting to believe that all impurity can be reduced to this kind of situation. However, we quickly hit a snag if we try to do this for other more general maps than states.

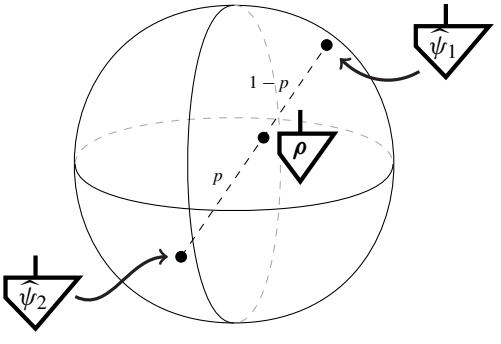
Theorem 6.73 Not every causal quantum map can be regarded as a mixture of pure causal quantum maps.

Proof Discarding cannot be decomposed in pure causal quantum effects, since there aren't any (cf. Theorem 6.54). □

The idea of mixing is not just important as a conceptual interpretation of impure quantum states (and some impure quantum maps), but it also provides a geometric picture for those states. In (6.44), $\boldsymbol{\rho}$ is a convex combination of pure states. The natural way to picture this is to think of $\boldsymbol{\rho}$ as lying somewhere between the pure states, where each p^i determines just how close it is to the i-th pure state (1 := 'at the same place', 0 := 'as far away as possible'). Recall from Section 6.1.2 that we can picture causal (i.e. normalised) pure quantum states in $\mathbb{C}^2$ as living on the surface of a sphere called the Bloch sphere. If we include arbitrary causal quantum states, we will include all convex combinations of states on the surface of the sphere. Therefore, we get not just a sphere, but a whole ball, called the *Bloch ball*. In this ball, a mixed state:

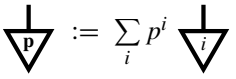


is pictured as a point inside the sphere:



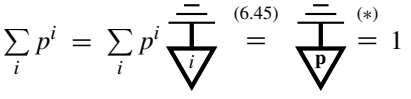
A special case of mixing is mixing over an ONB, which enables us to encode any probability distribution.

Proposition 6.74 Given a fixed ONB, probability distributions can be equivalently represented as causal quantum states of the form:



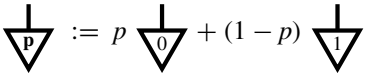
(6.45)

Proof We need to prove that the numbers p^i are all positive, which follows from Corollary 5.72, and that causality (cf. $(*)$ below) forces the probabilities to sum up to one:

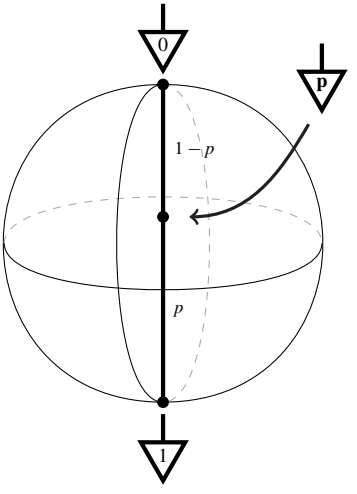


□

In the two-dimensional case probability distributions then become:

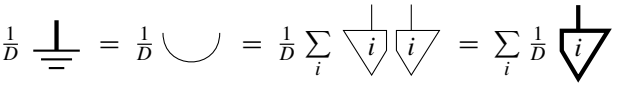


that is, they depend only on the number $p \in [0, 1]$, and we can visualise them within the Bloch ball as a line connecting the two doubled basis states:



If this were classical probability theory, that would be the end of the story: any probabilistic state decomposes uniquely into a probability distribution over the pure states (i.e. those corresponding to point distributions) spanning the space of probability distributions. However, for quantum states, the situation is quite different, in that this decomposition into pure states is usually not unique: a quantum state ρ may decompose as many different mixtures of pure states. The most extreme example of this phenomenon

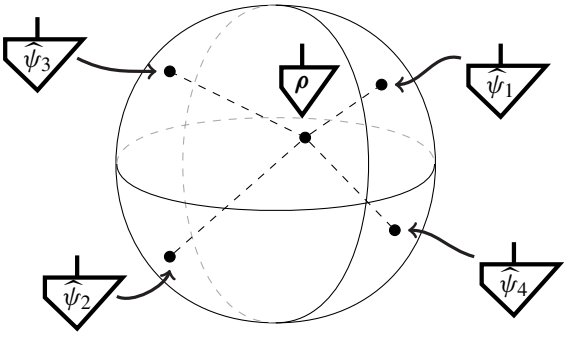
is the maximally mixed state, which can be seen as an equal mixture of the pure states corresponding to any orthonormal basis. Indeed, using Proposition 5.56, we can decompose the cup across any ONB:



(6.46)

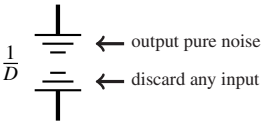
This also means that the maximally mixed state is 'equally distant' from any pure state, which explains the name 'maximally mixed': it has no bias towards any of the pure states.

So, given a mixed state, there is no unique interpretation in terms of a set of pure states. In fact, there is no particular reason to decompose $\boldsymbol{\rho}$ as a mixture over a basis or as a mixture of just two pure states. Why not three, or four:

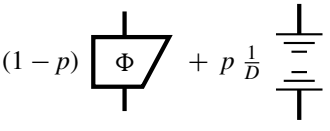


or a billion? In fact, mixed states are often used to describe enormous numbers of quantum systems, called *ensembles*, in a variety of pure states. Such an ensemble could represent, for example, all the photons in a laser beam.

Example 6.75 (noise) One can think of the maximally mixed state as total noise, in that it doesn't have any bias towards any meaningful data, i.e. any pure state. Consider the following process:



This process converts any input into noise. The fact that nothing from the input remains at the output is clear from the diagram in that it is ∘-separable. We can now mix this *noise map* with any other process:

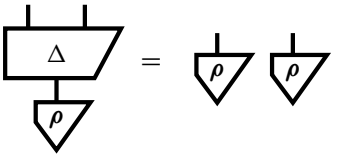


to make it a bit noisy. The bigger p is, the more noisy it is.

This is all we say about mixing for now. Section 8.3.4, where we picturalise mixing, contains diagrammatic counterparts of some of the results we have seen here, as well as some new results. In Chapter 13 we will introduce a *resource theory* that will let us compare and quantify mixedness.

6.2.8 The No-Broadcasting Theorem

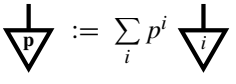
We already saw in Section 4.4.2 that any process theory that admits string diagrams satisfies a no-cloning theorem. Thus, in **quantum maps** there exists no quantum map Δ that clones all input states ρ:



Even if we restrict ourselves to pure states $\widehat{\psi}$, then there still do not exist cloning maps, since pure quantum maps already admit string diagrams. Restricting to causal states doesn't help either, since non-causal states can always be made causal by multiplying by a number.

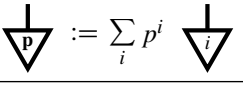
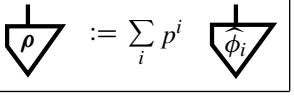
Section 4.4, where we first introduced no-cloning, was entitled 'Quantum Features from String Diagrams'. But how justified was this title? When one says 'quantum feature' one typically doesn't just mean something that happens to be true for quantum theory, but moreover, something that fails to hold classically.

In the previous section, we saw that we can represent a probability distribution as a quantum state as follows:



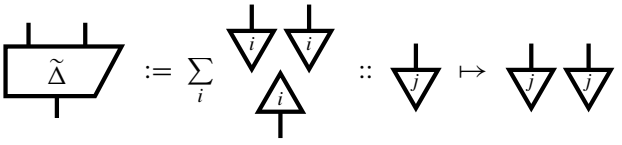
The pure states are then given by point distributions:

while all other probability distributions are mixtures of point distributions. This results in the following analogy:

	pure	mixed
probability distributions		
causal quantum states		

If no-cloning is a truly quantum feature, then it should be the case that we can clone classical states, i.e. probability distributions.

The good news is that there does exist a cloning map for point distributions, which we can represent as a quantum map:



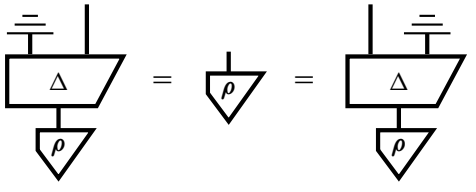
(6.47)

so while we cannot clone pure quantum states, we can clone pure probability distributions. However, the bad news is that this does not extend to general (i.e. mixed) probability distributions:

(6.48)

Just as in the quantum case, there is no map that clones all probability distributions. So 'no-cloning' only separates pure classical and quantum theories. What we really want is a weaker criterion, which still holds in the (mixed) classical world but fails for quantum maps.

This can be realised as follows. Rather than asking for the existence of a cloning map, we can ask for the existence of a *broadcasting map*, that is, a map that takes in a state ρ and outputs two systems with the property that if we discard either system, the overall state will be ρ:



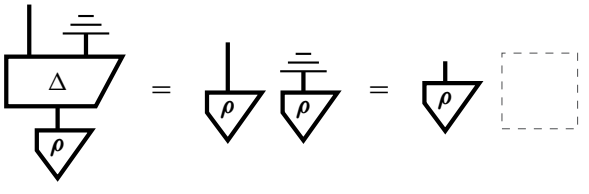
(6.49)

Here the term 'broadcasting' is used in the same sense as a television broadcast. When you receive a TV show, you only really care that it comes to you correctly, and not about what's going on everywhere else.

First we show that broadcasting is indeed weaker than cloning.

Proposition 6.76 Any cloning map is also a broadcasting map.

Proof For any cloning map Δ we have:



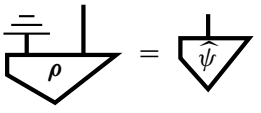
and the second equation also holds by symmetry. □

The converse is not true. In particular, we can broadcast probability distributions.

Exercise 6.77 Show that $\widetilde{\Delta}$ is a broadcasting map for probability distributions.

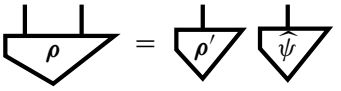
However, despite the fact that broadcasting is weaker than cloning, we still cannot broadcast quantum states. Thus no-broadcasting is a truly quantum feature, which doesn't depend on any extra assumptions such as purity. To prove no-broadcasting for quantum states, we first need to generalise the bipartite state in Proposition 6.33 from a pure state to any quantum state.

Proposition 6.78 If the reduced state of any quantum state is pure:



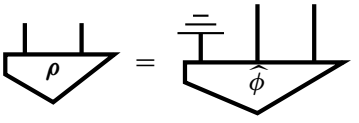
(6.50)

then the state ρ separates as follows:

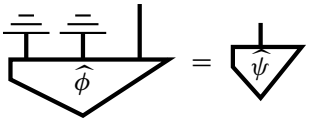


for some (causal) quantum state ρ'.

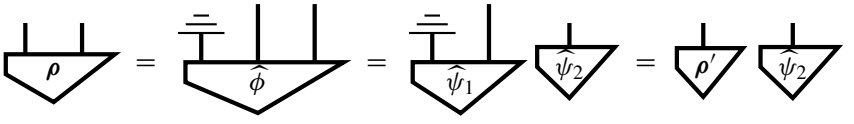
Proof Assuming (6.50), we first purify ρ (cf. Definition 6.47):



and substitute this into (6.50):



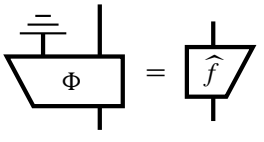
By Proposition 6.33, $\widehat{\phi}$ is $\otimes$-separable, in which case:



Just by multiplying $\widehat{\psi}_2$ by some number, we can always choose ρ' to be causal. In that case, equation (6.50) implies that $\widehat{\psi}_2 = \widehat{\psi}$. The proof of the converse is straightforward. □

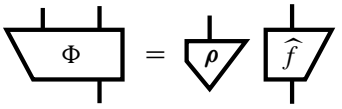
... and then we generalise to processes.

Proposition 6.79 If the *reduced map* of any quantum map Φ is pure:



(6.51)

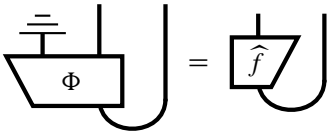
then Φ separates as follows:



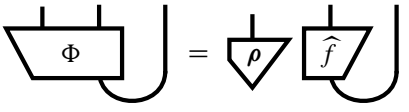
(6.52)

for some (causal) quantum state ρ.

Proof Bend the wire in (6.51):



Since the reduced state of the tripartite quantum state is pure, by Proposition 6.78 it separates as follows:

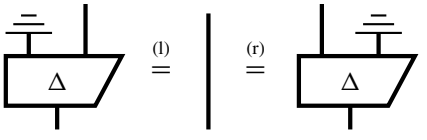


Unbend the wire and we're done. □

Now we are ready to prove the no-broadcasting theorem.

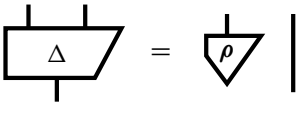
Theorem 6.80 Quantum states cannot be broadcast.

Proof By Theorem 6.24 there exists a basis consisting of quantum states for any type, so equations (6.49) are equivalent to:



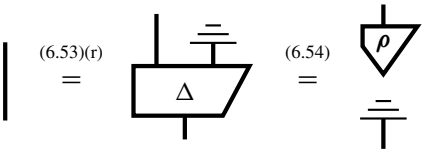
(6.53)

We now show there exists no such Δ. By equation (6.53) (l) the reduced state of Δ is pure, so by Proposition 6.79 we have:



(6.54)

for some state ρ. Hence it follows that:



Since the identity is ∘-separable, so is every other process involving that type, and hence the system must be trivial for Δ to exist. □

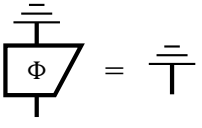
Remark 6.81 Even though we posed Theorem 6.80 specifically for **quantum maps**, we can derive this result for many process theories, much as we did for the no-cloning theorems in Section 4.4.2. Indeed, defining broadcasting directly via equation (6.53), no-broadcasting holds for the doubled version of any (non-trivial) process theory satisfying (5.47). Alternatively, if we are working in a process theory that didn't arise from doubling, we can define 'purity' of quantum processes as equation (6.51), implying equation (6.52). Then, no-broadcasting holds for any system whose identity process is 'pure' in the sense of (6.51). There is, however, a caveat with this approach, which we reveal in Section 6.7. Can you guess what it is?

6.3 Relativity in Process Theories

By now it must be clear that causality plays a central role in quantum theory. For example, it forces all pure quantum maps to be isometries and general quantum maps to arise as a purification of an isometry via Stinespring dilation. We motivated causality with this motto:

If the output of a process is discarded, it may as well have never happened.

This is embodied in the equation:



(6.55)

This also means that if a process is happening somewhere else, and its output never reaches us, we don't need to care about it. As we already noted, this is crucial to being able to even do science, in that it allows us to safely ignore parts of the universe that won't affect us.

The goal of this section is to show that the causality postulate guarantees compliance of quantum theory with the theory of relativity. In particular, causality guarantees that:

Nothing can attain speeds that are faster than light.

This is perhaps the most unexpected, and compelling, of its interpretations. In fact, we shall also see that quantum picturalism already has relativistic aspects built in and that causality also has the fact that the laws of physics should be observer independent built in.

6.3.1 Causal Structure

The theory of relativity teaches us about the structure of space and time, or in short, *spacetime*. Rather than there being a single theatre in which things take place with a single clock ticking, Aleks and Bob each have their own private screening of reality as well as

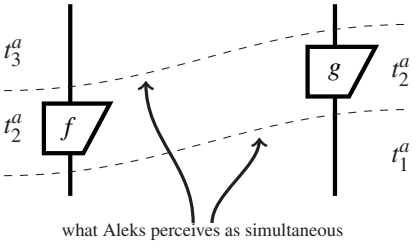
their own personal perception of how fast things progress. Relativity in its simplest form, called *special relativity*, is extracted from two principles:

1. all observers experience the same laws of physics, and
2. the speed of light through empty space is constant.

From these principles and a healthy dose of creativity one can then derive the entire theory of relativity, including the following features:

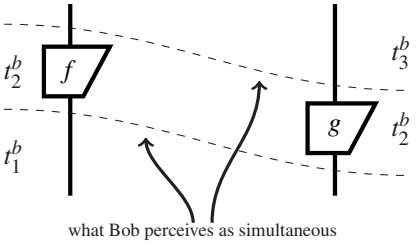
- *Relativity of simultaneity*, that is, two processes that appear to Aleks as happening at the same time, might appear to Bob as happening at very different times, and vice versa.
- *No faster-than-light travel*, that is, no object can ever attain speeds beyond the speed of light, and can only achieve that speed if it has no mass, since otherwise this would require an infinite amount of energy.

The first of these features implies that this scenario:



(6.56)

is equal to this scenario:

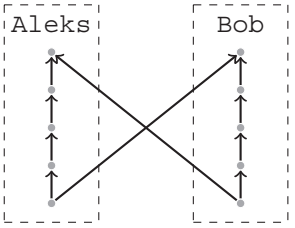


(6.57)

because in each case, for at least one observer these processes appear as simultaneous. Of course, we have assumed this right from the start, since those diagrams are equal! So by using the diagrammatic language we account for this important aspect of the theory of relativity.

The notion of simultaneity, which has lost its absolute (i.e. objective) meaning, makes way for the weaker (but still objective) notion of *spatial separation*, which allows for processes to appear as simultaneous to some, but not to others. We can depict this situation as two points without a connection between them:

Since no signal from Aleks to Bob can exceed the speed of light, when they are very far apart it takes some time before anything sent by Aleks will reach Bob, and vice versa. To visualise this, we can extend the picture above:



(6.58)

The points represent processes, performed at a certain location in spacetime, and the arrows indicate the possibility of one process to effect another.

We see in (6.58) that Aleks and Bob are free to perform a number of processes locally, but Aleks' processes can't have an effect on Bob's (and vice versa) until some time has passed after they are performed. Note how we have omitted such details as 'how far' Aleks and Bob are apart, or 'how long' each of the processes take. Such a drawing of points and arrows is what is called a *causal structure*.

The set of all points reachable by following the arrows from a single point is called the *causal future* of that point. One should think of these as all of the spacetime points one could reach from a given point by travelling less than or equal to the speed of light. In relativity parlance, this is called the *future light cone*. Similarly, there is a *past light cone*, and together these form the *light cone* for that point.

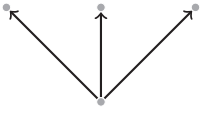
There are of course infinitely many causal structures one could imagine to describe processes interacting across space and time. Essentially, the only limit on what makes a 'valid' causal structure is that one (usually) assumes it does not include any directed cycles, much as we did in the case of circuit diagrams in Section 3.2.3.

Remark* 6.82 One can also define a causal structure as a *partially ordered set*. That is, a set with a relation $\leq$ where, for all a, b, c:

- $a \leq a$,
- $(a \leq b \text{ and } b \leq c) \implies a \leq c$, and
- $(a \leq b \text{ and } b \leq a) \implies a = b$

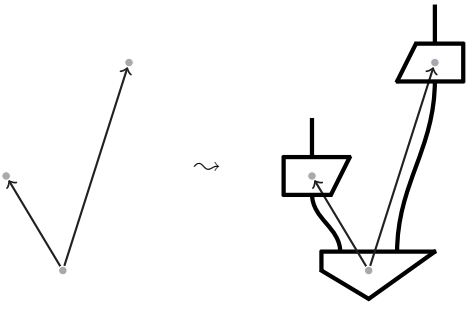
There is a tight relationship between the directed acyclic graphs derived from circuits (cf. Remark* 3.21) and causal structures, in that each directed acyclic graph can be turned into a partially ordered set by taking the transitive closure. Hence, in Remark* 6.82 we were right to characterise circuits as those diagrams that admit a causal structure.

Examples of causal structures are the V-shape and pitchfork-shape:



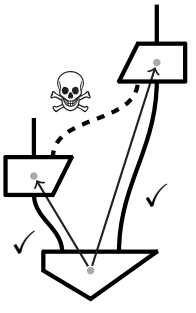
We will encounter the pitchfork-shape in Section 11.1 when we prove that quantum theory is *non-local*, and the V-shape will help us establish the connection between the causality postulate and the theory of relativity.

Assume now that we have some fixed causal structure that represents spacetime and a process theory that tells us which processes can take place (e.g. causal quantum maps). We can now think of those processes as happening at points in spacetime by laying them out on top of the causal structure:



(6.59)

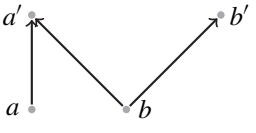
Notably, the arrows in the causal structure tell us where wires are allowed and where they are forbidden:



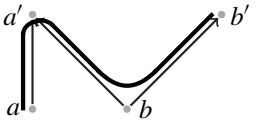
(6.60)

A process theory is said to be *non-signalling* if each process Ψ in a diagram with a fixed causal structure can have an influence only on processes in the causal future of Ψ. Naturally, a theory where processes can have effects on those outside their causal future is called *signalling*.

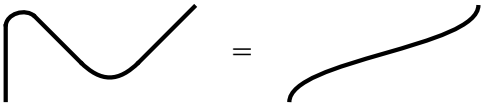
This is a non-trivial requirement of a process theory. In other words, there exist theories where processes can have an effect on other processes not in their causal future, even when we restrict to diagrams that respect the causal structure, as in (6.59). To see this, consider the following N-shaped causal structure:



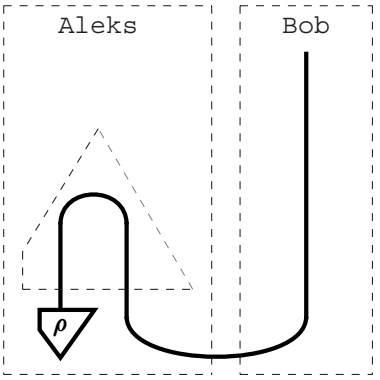
and any process theory that admits string diagrams (e.g. possibly non-causal **quantum maps**). We can clearly violate the causal structure by introducing a cup at b and a cap at a':



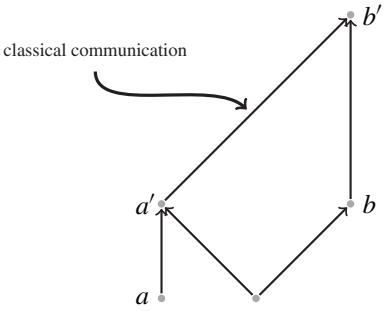
While in the causal structure there is no edge from a to b', one can still send data from a to b since we have:



In fact, this is exactly what happens in quantum teleportation:



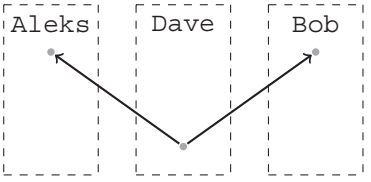
So, what's going on here? Is quantum theory signalling because it allows teleportation? Absolutely not! By just looking at a single 'branch' of the teleportation protocol, that is, a fixed error $\widehat{U}_i$, we are overlooking an important part: the communication of Aleks' value i to Bob, which is necessary to correct that error. If we account for this classical communication, the causal structure is again respected:



We will see in Section 6.4.4 that if Aleks does not send Bob his value i, the state Bob gets at the end will just be noise, with no trace of ρ whatsoever.

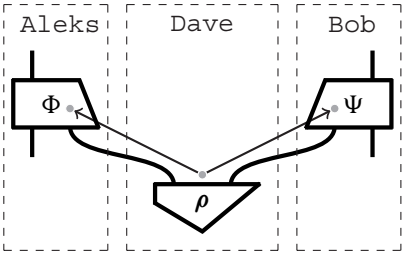
6.3.2 Causality Implies Non-signalling

To show this, we will rely on the simple V-shaped causal structure:



That is, Aleks and Bob may have some shared history, perhaps when Dave gave them two halves of an entangled state. However, now they have moved far away from each other, so far in fact that they can no longer communicate with each other without sending messages faster than the speed of light, which, of course, they can't do.

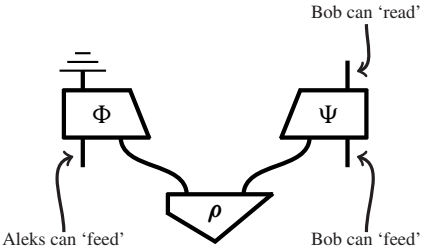
We can decorate the causal structure with a diagram of processes:



When we discussed teleportation before, we considered very specific processes, but now we will merely assume that all of the processes obey causality. In fact, the following argument goes through even if Φ, Ψ and $\boldsymbol{\rho}$ are not quantum processes, as long as they still satisfy the causality equation with respect to the discarding processes.

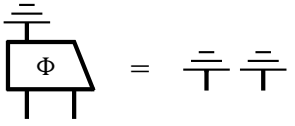
The causal structure allows Aleks and Bob to have local inputs and outputs, so that they can control their processes and receive data from them. Technically, it allows inputs/outputs for Dave as well, but we won't need them. But crucially, since there is no arrow in the causal structure from Aleks to Bob or vice versa, it should not be possible for either of them to send a signal to the other. This means that neither of them should be allowed to derive something about the other one's input using just their own inputs and outputs. Otherwise, this could be exploited to communicate some data, i.e. signal, in a way that violates the causal structure.

Our claim is that non-signalling follows from nothing but the causality postulate (6.55). We start by considering the case of Bob learning something about Aleks' input. From Bob's perspective, we should discard Aleks' output since it is not accessible to Bob:

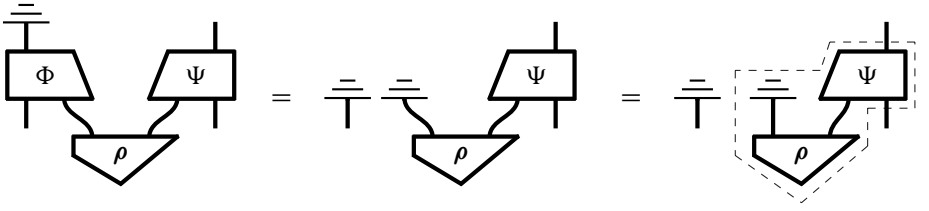


(6.61)

Let's now see if Bob can learn anything about Aleks' input from his own input-output pair. By causality we have:



and hence it follows that:



So from Bob's perspective, his input-output pair is disconnected from Aleks' input, which is discarded. Thus no signalling from Aleks to Bob can take place. By symmetry it also follows that Bob cannot signal to Aleks. Thus we have the following.

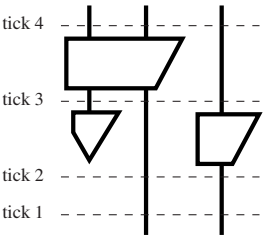
Theorem 6.83 If a process theory has a discarding process for each type and it satisfies causality, then it is non-signalling.

Exercise 6.84 Above we only established non-signalling for the V-shaped causal structure. What if we extend this to a diamond-shaped one, that is, we add a point in the joint future of Aleks and Bob? More generally, show that non-signalling holds for any causal structure.

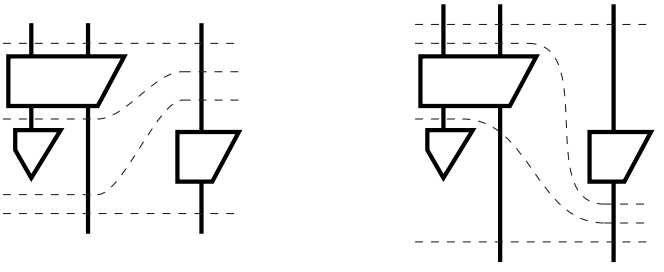
6.3.3 Causality and Covariance

We will now show that distinct observers Aleks and Bob will always experience the same laws of physics. More precisely, we will show that states restricted to a single location in spacetime (known as *local states*) look the same to both. Consequently, whatever laws of physics made the system end up in that state must be the same too.

So, what does it mean to ask what a local state 'looks like' for a given observer? Recall from Section 6.3.1 that the notion of simultaneity depends on the observer. Suppose we chop up a diagram into a series of layers indicating what processes have happened by the time a clock ticks for a single observer:

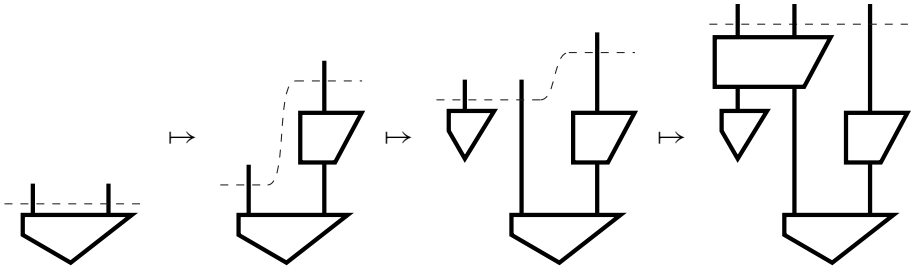


This way of chopping up a diagram is called a *foliation*, and since the notion of 'simultaneous' depends on the observer, so too do foliations. Thus a single diagram admits many different ones:

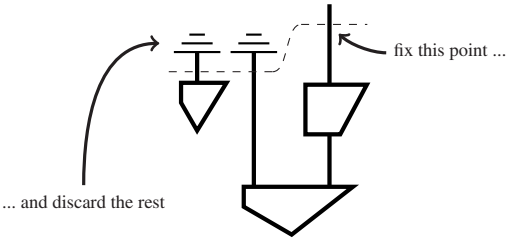


Another way of saying the laws of physics are the same for all observers is to say they are foliation-independent, or *covariant*.

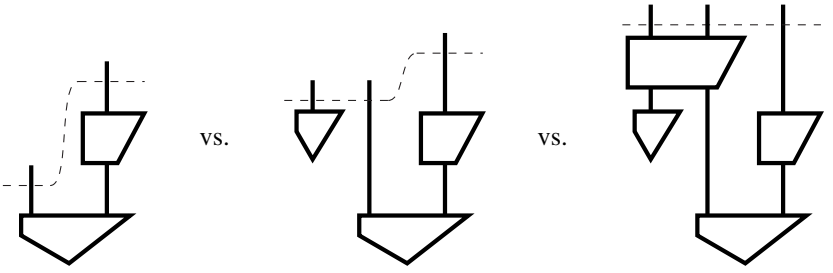
If we input an initial state into a foliated diagram, we can compute the resulting state at each layer, so we can see it evolving in time according to a particular foliation. For, example the left foliation above yields:



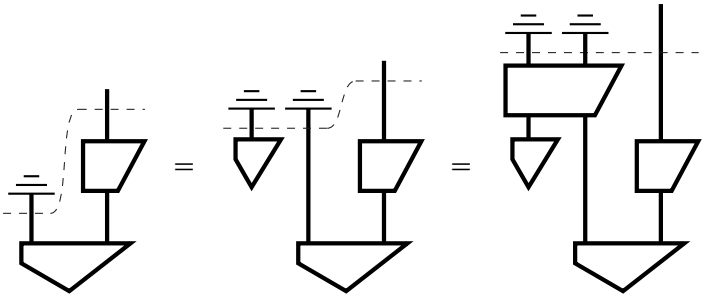
A local state is obtained by fixing a single spacetime point, and discarding all of the other systems in a single layer, e.g.



Of course, there are typically many different layers that include a single spacetime point, coming from different possible foliations:



But the causality postulate guarantees that the local state does not depend on the choice of layer:



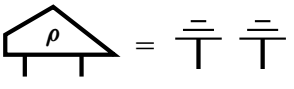
Hence it also does not depend on the choice of foliation.

6.4 Quantum Processes

In the introduction to this chapter we announced that causality will be a defining constraint on general quantum processes. These quantum processes can have more than one quantum map, which, as a whole, obey a version of the causality postulate that generalises the one for quantum maps. A single quantum map satisfying causality is then a special case, namely, that of a deterministic quantum process. However, deterministic quantum processes are scarce. For example, by Theorem 6.54 there is only one causal quantum effect for any type: discarding. The true potential of quantum processes only becomes apparent when also considering the non-deterministic ones. Without these, things like teleportation are simply not possible, since the cap effect that we relied on to perform teleportation is not causal, nor is any non-separable bipartite effect.

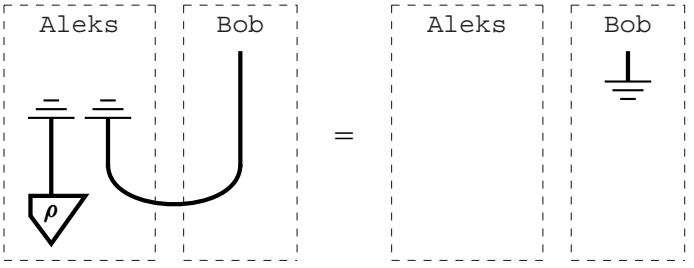
Proposition 6.85 All causal quantum effects are separable.

Proof Since by equation (6.16) discarding multiple systems is the same as discarding each system individually, from Theorem 6.54 it follows that:



which, in particular, is separable. □

Now, suppose we try to do quantum teleportation with the only causal bipartite quantum effect available:



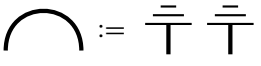
(6.62)

The best that Bob can obtain is the maximally mixed state, that is, nothing but noise. In particular, there is absolutely no trace left of the state ρ.

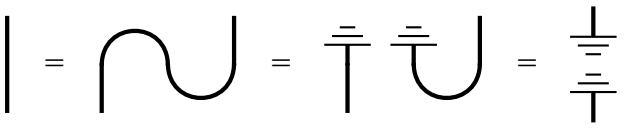
The fact that we don't have caps around is bad news, as indicated by the following clanger.

Proposition 6.86 Circuits of causal quantum maps are again causal quantum maps, so causal quantum maps form a process theory. However, that theory does not admit string diagrams.

Proof The proof that circuits of quantum maps are again quantum maps is included in the proof of Proposition 6.89 below. Now, suppose causal quantum maps did admit string diagrams. Then, by Proposition 6.85:



from which it follows that the identity is disconnected:

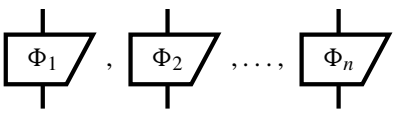


and hence the process theory must be trivial. Since the process theory of causal quantum maps is not trivial, it cannot admit string diagrams. □

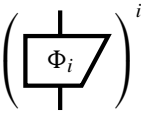
Fortunately, by allowing for non-determinism, we will be able to reintroduce all that was lost, and hence rescue string diagrams!

6.4.1 Non-deterministic Quantum Processes

An (uncontrolled) *quantum process* is a collection of quantum maps:

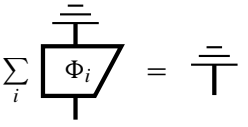


which we shall denote:



(6.63)

that together satisfy the *causality postulate*:

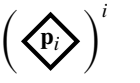


(6.64)

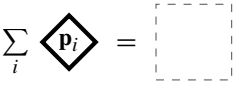
We refer to the elements in the set (6.63) as *branches*. If there is only one branch, then we call the quantum process *deterministic*; otherwise, we call it *non-deterministic*. When a system undergoes a quantum process, one of the branches actually happens. We call this branch (or simply its index i) the *outcome* of the process.

Remark 6.87 We say 'uncontrolled' quantum process because in Section 6.4.5, we will generalise this definition to allow a quantum process to depend on (i.e. 'be controlled by') the outcome of an earlier process.

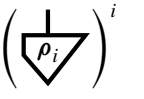
Physics is all about making predictions, so if many alternatives are possible, one wants to know how probable each of these alternatives is. This is precisely what quantum processes consisting of numbers:



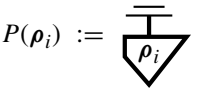
do for us. We already know that doubling guarantees that the numbers are positive, and this new multibranch version of causality guarantees that these numbers moreover form a probability distribution:



We can also associate a probability distribution to quantum processes consisting of states. Given such a quantum process:

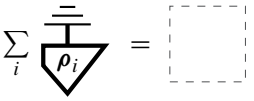


the weight of each state provides its probability:

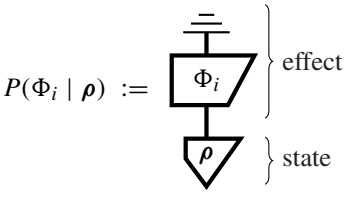


(6.65)

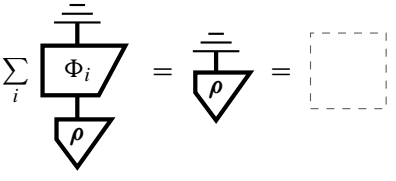
As mentioned in Section 6.2.3, this is nothing but an instance of the Born rule. Again, causality means that these probabilities add up to one:



Unlike states, we cannot associate a fixed probability distribution to general quantum processes. This should come as no surprise, since in the case that a quantum process is applied to a system in a certain state, the probabilities may of course depend on that state. For example, although we haven't yet said what a quantum measurement is, it should be evident that the outcome of performing anything called a 'measurement' will depend on what is actually being measured. However, once we apply a quantum process to a state, we can again assign probabilities, and the Born rule tells us what they are:



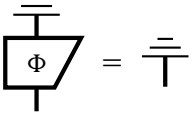
and once again causality guarantees that they add up to one:



Thus, the causality postulate (6.64) for a generic process guarantees that:

probabilities can be consistently assigned to branches.

In the case of deterministic processes, causality reduces to the form in which we already encountered it in Definition 6.52:

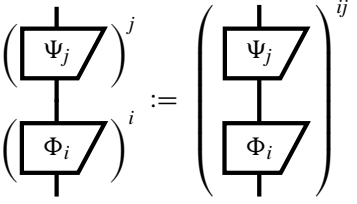


So quantum processes generalise causal quantum maps, and in particular, a deterministic quantum process consists of a single, causal quantum map.

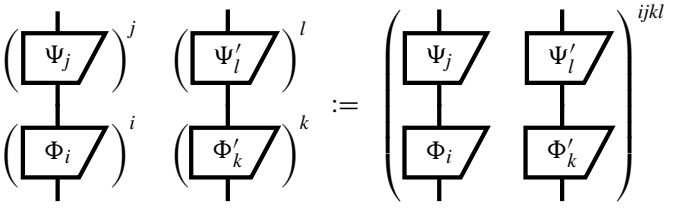
Remark 6.88 In the form of (6.64), the causality postulate doesn't seem to admit the interpretation we had provided it with in the case of quantum states and quantum maps, namely, that discarding the outputs of a process is the same as discarding its inputs without the process ever happening. This is because knowledge of which branch Ψ_i happened should actually be treated as a *classical output* of the process. Once we have a diagrammatic

account on classical data, it will become clear that summing together all of the branches is indeed the result of 'discarding' this classical output, in which case (6.64) says the process might as well have never happened.

Now, consider what happens if we compose two non-deterministic processes sequentially. If the first process has, say, four branches (indexed by i), and we feed its output to another process that has three branches (indexed by j), then any combination of the i-braches and the j-branches could happen. So, the resulting process:



has $4 \cdot 3 = 12$ branches (indexed by ij). We can use a similar rule to compose processes in parallel, and even make arbitrary circuits:

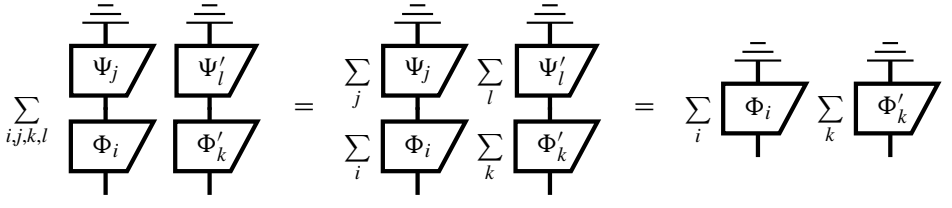


(6.66)

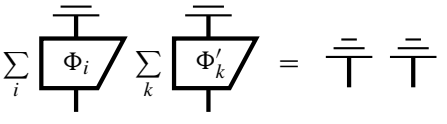
That is, we can pull these braces to the outside of the picture, much like we could do with sums, which, as we will see in Chapter 8, is not entirely a coincidence. The result is again a quantum process because of the following.

Proposition 6.89 Causality is preserved when forming circuits of quantum processes.

Proof We show this for the composition structure in (6.66), which encompasses both parallel and sequential composition, and hence arbitrary circuits. By causality of the quantum processes $(\Psi_j)^j$ and $(\Psi'_l)^l$ we have:



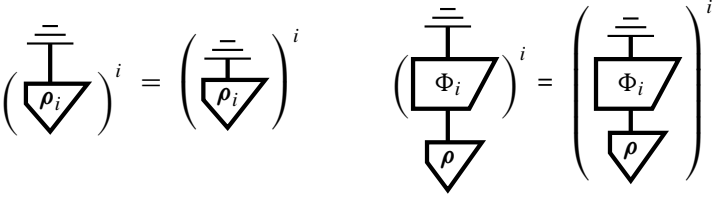
and by causality of the quantum processes $(\Phi_i)^i$ and $(\Phi'_k)^k$ we have:



that is, causality of the four quantum processes in the LHS of (6.66) implies causality of the RHS of (6.66). □

Remark 6.90 Quantum processes 'almost' form a process theory. When we defined process theories back in Chapter 3, we said they give an interpretation for diagrams of processes, not diagrams with these funny parentheses $(\ )^i$ in them. We'll fix this in Chapter 8 when we replace parentheses with classical wires and get a bonafide process theory. But in the meantime, there's no harm in treating this 'almost' process theory as if it really were one.

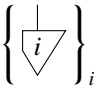
Example 6.91 Thanks to composition of quantum processes, the three cases we gave earlier of assigning probabilities to branches all boil down to the first case of quantum processes consisting of numbers:



6.4.2 Non-deterministic Realisation of All Quantum Maps

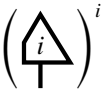
Since non-deterministic quantum processes are going to rescue our beloved string diagrams, then at the very least they should provide us with some effects other than discarding. The following is an example of a quantum process that does just that.

Proposition 6.92 For any ONB:



(6.67)

the corresponding doubled effects form a quantum process:



(6.68)

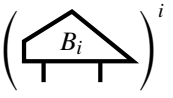
Proof Vertically reflecting (6.46) without the $\frac{1}{D}$-factor yields:

so (6.68) is indeed causal, and hence a quantum process. □

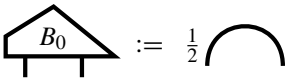
This is in fact one of the most important examples of a quantum process, which is called an *ONB measurement*. We will study this and related processes extensively in the next

chapter. For now, it will provide us with a sufficient dose of non-determinism to generate arbitrary quantum maps.

Since the Bell basis is an ONB (cf. Section 5.3.6), by Proposition 6.92 there is a corresponding quantum process:

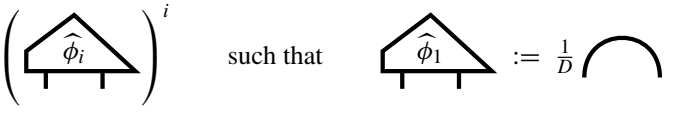


which up to a number includes the doubled cap as a branch:



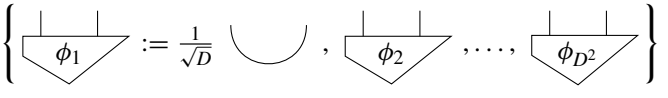
This holds not just for the cap for qubits, but also for arbitrary dimension, which, in turn, will allow us to recover all quantum maps.

Lemma 6.93 Bell effects can be realised non-deterministically. Explicitly, there exists a quantum process:



(6.69)

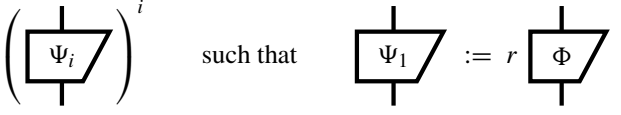
Proof Since by Proposition 5.79 any normalised state can be regarded as a member of an ONB, there exists an ONB containing the normalised cup:



(6.70)

Then, by Proposition 6.92 it follows that (6.69) is a quantum process. □

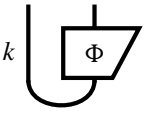
Theorem 6.94 Every quantum map Φ can be realised non-deterministically, up to a number. Explicitly, there exists a quantum process:



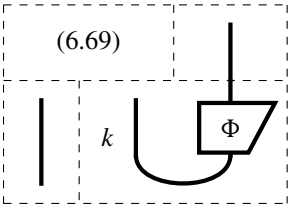
(6.71)

for some $r > 0$.

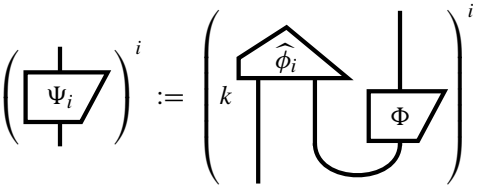
Proof Choose k such that the following state is causal:



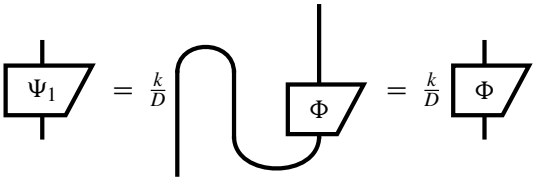
When we compose this quantum process with the quantum process (6.69) that we constructed in Lemma 6.93:



we obtain the quantum process:

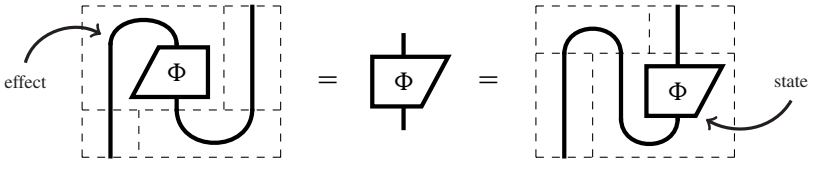


where:



which completes the proof, for $r := \frac{k}{D}$. □

Exercise 6.95 A quantum map can be realised by either encoding it as a state or an effect:

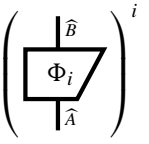


Construct an alternative proof of Theorem 6.94 by encoding Φ as an effect that occurs as a branch of a non-deterministic process.

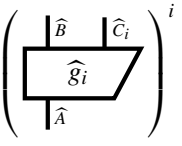
6.4.3 Purification of Quantum Processes

By Proposition 6.46 we know that each quantum map can be purified, and hence, all deterministic quantum processes can be purified. In fact, this was pretty much built in to the notion of a quantum map, which is defined to be the result of composing pure quantum maps with discarding. Purification has the appealing interpretation that every quantum map arises from a pure one by discarding an output, and, by causality, this pure quantum map can be assumed to be an isometry (Stinespring dilation).

At first glance, it seems to trivially follow that we can also purify non-deterministic quantum processes. Indeed, each branch in any quantum process:

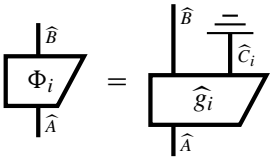


can be purified, that is, there exist:



(6.72)

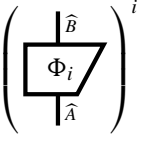
such that:



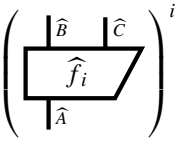
However, this is where we hit a bit of a snag. For each i the type $\widehat{C}_i$ might be different, in which case the processes $\widehat{g}_i$ may have different output types, meaning (6.72) is no longer a well-defined quantum process.

Fortunately, this problem can be fixed by finding a suitable *joint purification* of the branches of a quantum process:

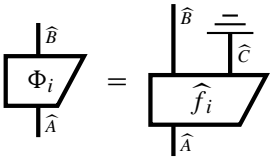
Lemma 6.96 For any quantum process:



there exists a single type $\widehat{C}$ and a quantum process:

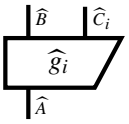


such that:

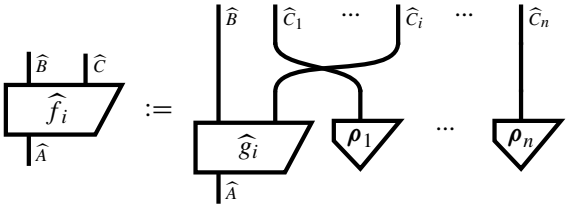


(6.73)

Proof We know that for each quantum map Φ_i there exists a purification:



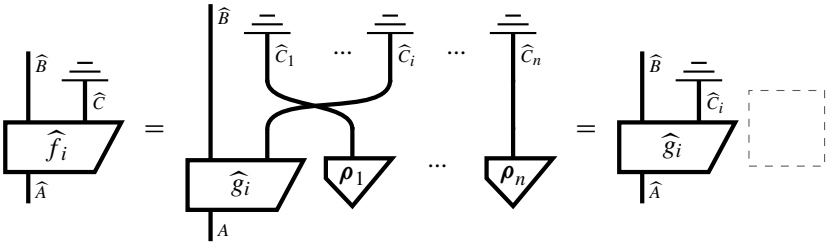
For every type C_j with $j \neq i$ pick a causal state $\boldsymbol{\rho}_i$. To each pure quantum map $\widehat{g}_i$ we then associate another pure quantum map $\widehat{f}_i$ as follows:



So in particular we have:

$$\widehat{C} := \widehat{C}_1 \otimes \cdots \otimes \widehat{C}_n$$

where n is the number of outcomes of the quantum process $(\Phi_i)^i$. By causality of all the states $\boldsymbol{\rho}_i$ we then have:

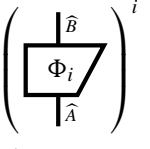


and hence (6.73) indeed holds. □

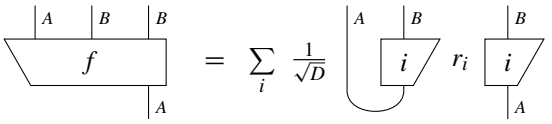
The choice of purification, and the choice of purifying system $\widehat{C}$ in particular, is not unique. While the proof of Lemma 6.96 gives a simple, diagrammatic way to construct $\widehat{C}$, it is not optimal in the sense that the dimension of C in general won't be the smallest possible one. In particular, when the maps Ψ_i are not already pure, the dimension of C from Lemma 6.73 grows exponentially with the number of maps.

On the other hand, we can stick an upper bound on how big C needs to be, which depends only on the dimensions of A and B. This upper-bound result is sometimes known as *Choi's theorem*.

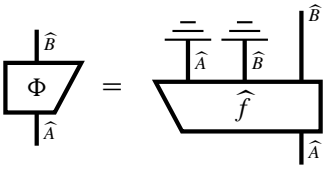
Exercise 6.97 Show that one can always purify a quantum process:



using the auxiliary system $\widehat{C}$ to be $\widehat{A} \otimes \widehat{B}$, by showing that the linear map of the form:

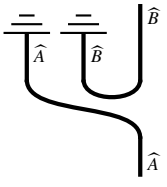


yields a purification of any quantum map Φ_i:



for a suitably chosen Hilbert–Schmidt basis. (Hint: rely on Corollary 6.69.)

Remark 6.98 While Exercise 6.97 gives an upper bound on the dimension needed for C, it can in general be smaller. For example, even applying Lemma 6.96 to a set of pure quantum maps will yield a smaller (in this case, trivial) system C. On the other hand, sometimes the full system $\widehat{A} \otimes \widehat{B}$ is necessary. For example, in:



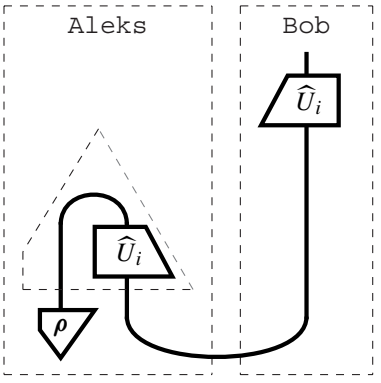
the auxiliary system $\widehat{C}$ clearly at least needs to be $\widehat{A} \otimes \widehat{B}$ for any joint purification procedure.

So far so good. We saved the idea of purification for the case of non-deterministic quantum processes. But what about Stinespring dilation? In particular, one may wonder if each branch could be chosen to be an isometry (possibly multiplied by some probability). The answer is no. One counter-example is any quantum process of effects (e.g. the ONB of effects from Proposition 6.92), since clearly no isometry can exist from a non-trivial system to a trivial one.

However, what we can do is associate a single isometry to any quantum process, which in fact, as we shall see in Section 7.3.4, does provide a very satisfactory generalisation of Stinespring dilation to non-deterministic quantum processes, called *Naimark dilation*.

6.4.4 Teleportation Needs Classical Communication

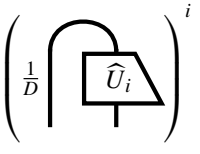
With what we've learned in this chapter, let's now try to fill in the quantum teleportation diagram:



with quantum processes.

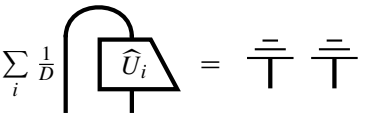
Starting with $\boldsymbol{\rho}$ and half of a Bell state, Aleks needs to perform some quantum process that produces a Bell effect, possibly with some error (represented by the pure quantum map $\widehat{U}_i$). Then, some time later Bob should perform another quantum process to correct the error.

We established in (6.62) that we definitely need non-deterministic quantum processes in order to realise teleportation, so Aleks should perform a non-deterministic process consisting of a set of effects:



(6.74)

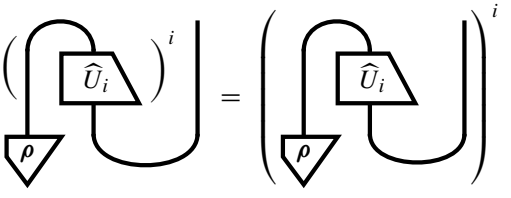
such that causality is satisfied:



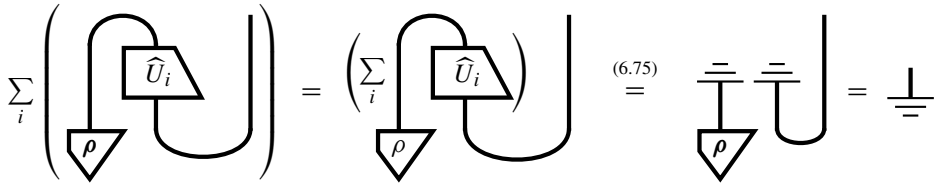
(6.75)

At this point Aleks communicates the value i produced by his quantum process to Bob and Bob performs his correction. Even though we know how to express non-deterministic quantum processes, we haven't yet seen an instance of a quantum process that depends on the outcome of another process. Before we go to the effort of defining such a thing, we should ask ourselves: is Bob's correction really necessary?

In other words, what happens if Aleks decides not to communicate the value i to Bob? Consider a version of teleportation without a correction:



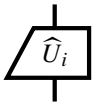
The overall process is a non-deterministic quantum state (we've omitted the $\frac{1}{D}$ for clarity). However, Bob doesn't have access to the value i. To account for this lack of information, what Bob perceives is actually just a mixture of each of the possible branches. Then, by relying on causality of the quantum process we obtain:



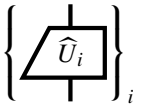
So Bob receives the maximally mixed state, with no sign of $\boldsymbol{\rho}$ anywhere. Hence we can conclude that without any classical communication, nothing can be teleported whatsoever.

6.4.5 Controlled Processes

Right, so instead of hiding the value i, Aleks must send it to Bob so that Bob can perform a correction depending on the outcome i:

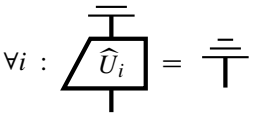


Now, while it might be tempting to think of the set:

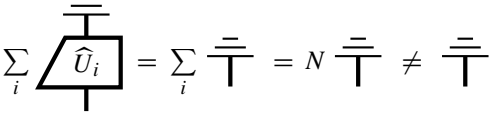


(6.76)

as a non-deterministic quantum process, this is not the case! In fact, since each of the elements is itself a quantum process:

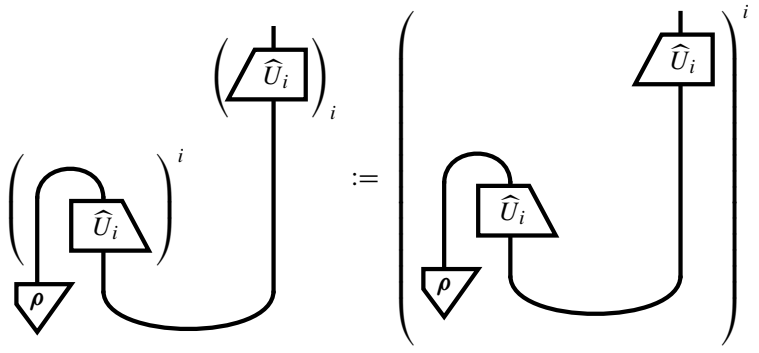


it follows the whole N-element set cannot be a quantum process:



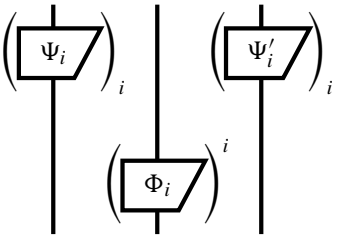
All together there is only one index i in play. The key difference between the set (6.76) and a quantum processes is that a quantum process produces a value of i, whereas (6.76) depends on it. In fact, there is no reason why we should restrict to deterministic processes here: depending on i we may wish to perform one (non-deterministic) quantum process out of a set of quantum processes. We will refer to such a set of quantum processes depending on some classical index as *controlled* by i.

So, the overall quantum teleportation process should look something like this:



where the brackets still indicate that there are several branches, indexed by i, but these branches are related to the index i in two different ways. The process with an upper index i produces the value of i, whereas the process with a lower index i depends on, or consumes, the value of i.

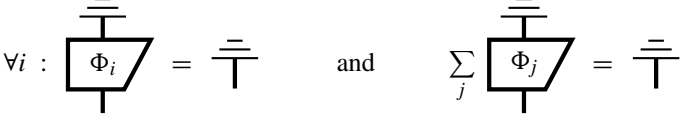
One can also imagine scenarios where i is consumed several times. For example, in:



the quantum process $(\Phi_i)^i$ produces i, while the two controlled quantum processes $(\Psi_i)_i$ and $\left(\Psi'_i\right)_i$ consume it. Although we can consume i any number of times, we cannot force several non-deterministic processes to have the same outcome, so i must be produced by exactly one process.

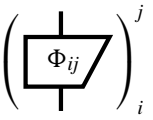
Remark 6.99 The use of upper and lower indices to model classical inputs/outputs mirrors that of the diagram formulas from Section 3.1.3. The fact that one classical outcome can be used multiple times comes from the fact that classical data can be *copied* (i.e. 'cloned'). We will see how this unique feature of classical data manifests in Chapter 8.

We are now ready to give a full definition of a quantum process, which includes the possibility of being controlled by a classical input. By combining the conditions:



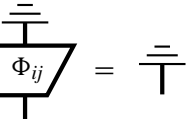
into one, we get the full definition of a quantum process.

Definition 6.100 A *quantum process* is a set of quantum maps:

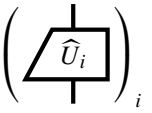


which satisfy:

$$\forall i: \quad \sum_j \quad = \quad \tag{6.77}$$



For $1 \leq i \leq m, 1 \leq j \leq n$, if $m = 1$ above, this definition yields our prior notion of an uncontrolled quantum process. If $n = 1$, this yields a controlled quantum process consisting entirely of causal quantum maps (i.e. without non-determinism), like Bob's unitary corrections:

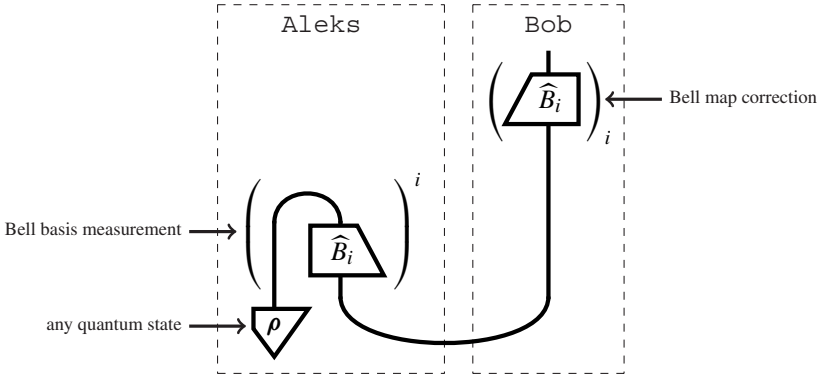


Of course, if both m and n are 1, this is just a single causal quantum map.

You may wonder why we only defined individual quantum processes rather than the theory of **quantum processes**. At this point the indices i for a quantum process are treated as something external to the diagram, rather than something that lives on a wire. This doesn't really fit in with our definition of a process theory, which is purely diagrammatic. We will fix this in Chapter 8 by replacing these indices with *classical wires*.

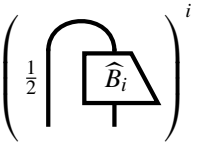
6.4.6 Quantum Teleportation in Detail

And finally, the moment you've been waiting for: we now have all of the ingredients to fully describe quantum teleportation. In the case of qubits, quantum teleportation is realised as follows:

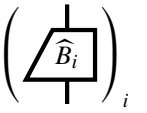


for B_i the ith Bell map. In words:

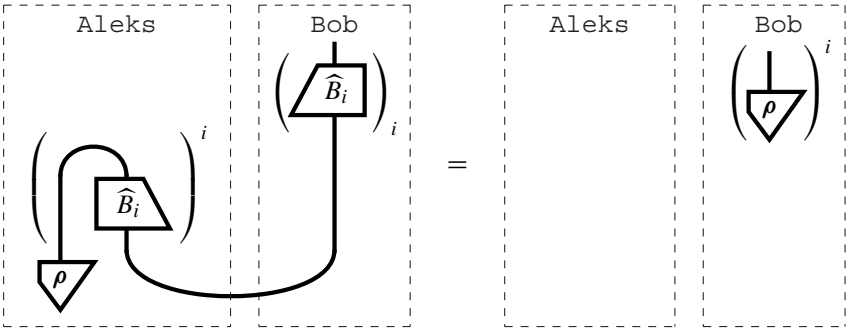
1. Aleks is in possession of an arbitrary quantum state ρ, which he wishes to send to Bob, and Aleks and Bob together share a Bell state.
2. Aleks performs a *Bell basis measurement*, that is, the following non-deterministic quantum process:



   on ρ and his half of the Bell state.
3. Aleks sends the outcome i to Bob.
4. Bob performs a *Bell map correction*, that is, the following controlled quantum process:



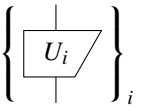
As we have seen many times by now, the result of performing this protocol is that, regardless of Aleks' measurement outcome, Bob will receive ρ:



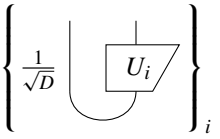
Back in Section 4.4.4 we also emphasised that teleportation does not enable one to magically beam something through space, but rather enables one to to use one kind of data (classical) to send another kind of data (quantum). Since there are four elements in the Bell basis, Aleks has to send a number from 0 to 3 (i.e. two classical bits) to Bob. In exchange, Bob receives Aleks' qubit, which could be any one of infinitely many points in the Bloch ball.

One remarkable thing about quantum teleportation is that, while it is essential that we rely on non-deterministic quantum processes, we produce a deterministic process overall. That is, no matter what outcome Aleks' measurement produces, the overall process will always give Bob the state ρ. The reason is of course that each of the branches, due to the correction made by Bob, gives rise to the same process. This shows that with clever tricks one can 'undo' the non-determinism of quantum processes.

While the protocol described just now is the simplest non-trivial case of a quantum teleportation protocol, there is nothing particularly special about the Bell matrices or qubits, in terms of the role they play in teleportation. To do teleportation, all we need is a set of unitaries:

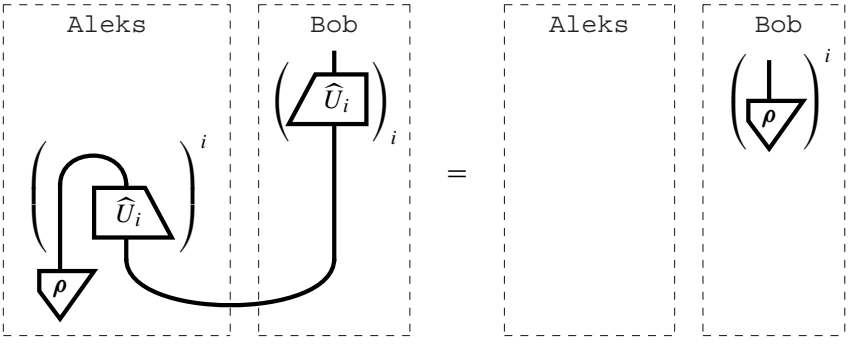


such that:

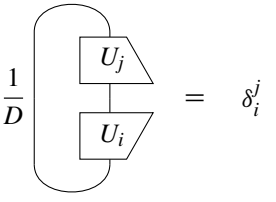


(6.78)

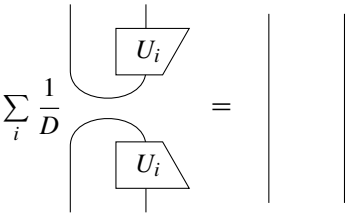
forms an ONB. Then, there is a process consisting of ONB effects and corresponding corrections such that we have:



Relying on Theorem 5.32, the conditions for the set (6.78) forming an ONB can equivalently be stated as the following two equations:



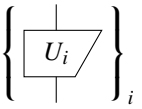
(6.79)



(6.80)

Thus we can conclude the following.

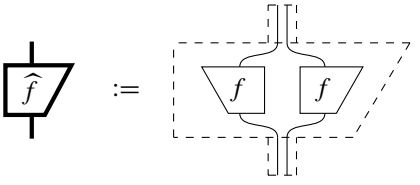
Corollary 6.101 A set of unitaries:



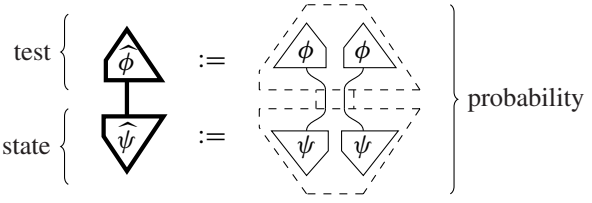
yields a teleportation protocol whenever (6.79) and (6.80) are satisfied.

6.5 Summary: What to Remember

1. **Pure quantum maps** arise from 'doubling' linear maps:

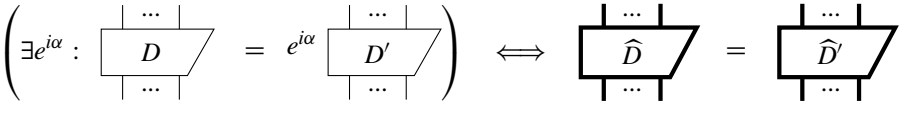


Composing pure quantum states and effects:



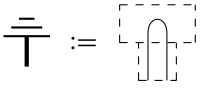
yields the *Born rule* for quantum theory. This process theory admits string diagrams, where the cups and caps are given by:

2. Let D and D' be arbitrary diagrams in **linear maps**, and $\widehat{D}$ and $\widehat{D'}$ be their doubled versions in **pure quantum maps**, then:

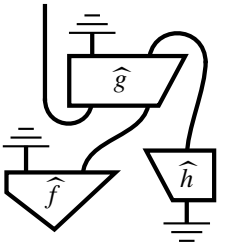


We refer to the number $e^{i\alpha}$ as a *global phase*.

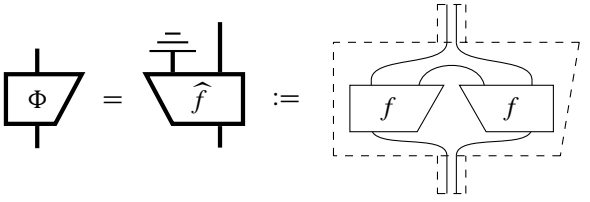
3. *Discarding* is the linear map:



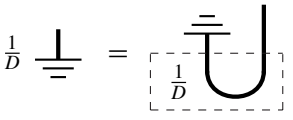
which fails to be a pure quantum map. The process theory of **quantum maps** arises from composing pure quantum maps with discarding:



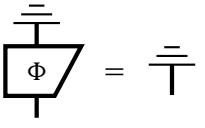
Any quantum map Φ can be written in the following form:



and we call $\widehat{f}$ a *purification* of Φ. For example, the *maximally mixed state* can be purified by means of the *Bell state*:



4. The *causality* postulate:



(6.81)

implements the following obvious (and necessary) fact:

If the output of a process is discarded, it may as well have never happened.

This implies that we can discard everything that is not 'connected' to what we care about, which is an obvious prerequisite for basic scientific practice. Causality also imposes compatibility with the theory of relativity, notably this fact:

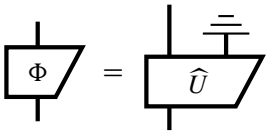
Nothing can attain speeds that are faster than light.

5. Causality imposes the following restrictions:

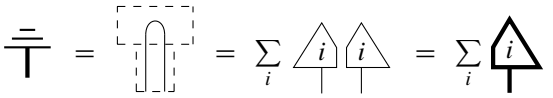
- all pure quantum maps are isometries, and hence
- all pure quantum maps from a system to itself are unitary.

The following fact is known as *Stinespring dilation*:

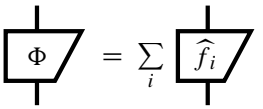
- Any causal quantum map Φ can be purified as an isometry $\widehat{U}$:



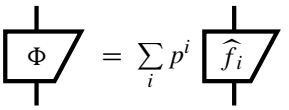
6. Since discarding decomposes for any ONB as follows:



every quantum map Φ admits a *Kraus decomposition*:



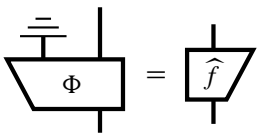
If such representation is moreover of the form:



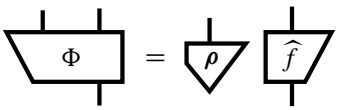
where $\{p^i\}_i$ is a probability distribution and all $\widehat{f}_i$ are causal, then we call it a *mixture*. Every causal quantum state can be regarded as a mixture of pure causal quantum states. However, this is not the case for every causal quantum map.

> ⚠ The notation for mixing involving sums is only a temporary one. In Chapter 8, we will picturalise these.

7. If the *reduced process* of any quantum process is pure:

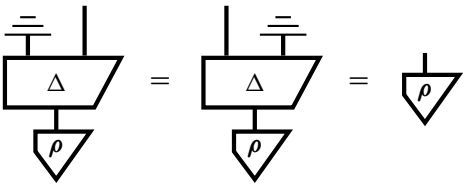


then the state Φ separates as follows:



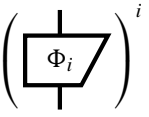
for some (causal) quantum state ρ.

8. Quantum states cannot be broadcast; i.e. there exists no quantum map Δ such that for all quantum states ρ we have:

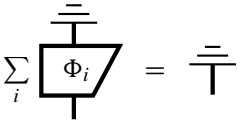


The key ingredient for proving this is **7**.

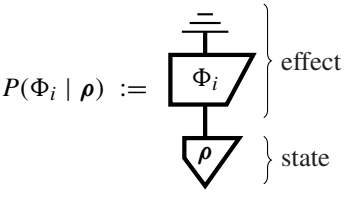
9. **Quantum processes** are lists of quantum maps:



satisfying (generalised) *causality*:

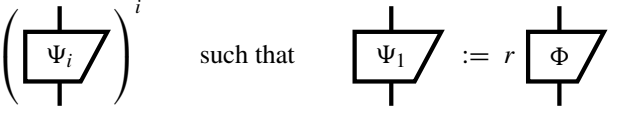


The quantum maps Φ_i are the *branches*, and if there is more than one branch, then the quantum process is *non-deterministic*. When a quantum process is applied to a state, then the probability of each branch is computed using the Born rule:



> ⚠ The notation involving braces and sums is only a temporary one. In Chapter 8, we will picturalise these too.

10. All quantum maps arise as branches of quantum processes. That is, for every quantum map Ψ, there exists a quantum process:



for some $r \geq 0$.

6.6 Advanced Material*

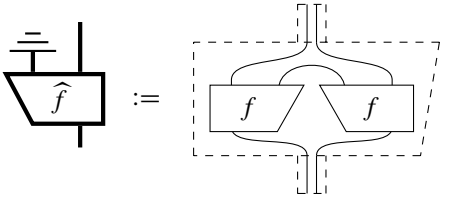
By the time we started this chapter, we had already made a huge, (mostly) unmotivated assumption: that quantum processes should be built up from **linear maps**. In this section,

we'll look at how one might go about getting around this assumption. In particular, we take a closer look at the *doubling construction* (which throughout this section includes adjoining discarding), and in particular, how it can be understood not just as a construction on **linear maps**, but something that can be applied to any process theory admitting string diagrams. By doing so, we can hope to understand or even reproduce all of the predictions of quantum theory in a way that is completely independent of Hilbert spaces and linear maps.

We end this chapter with something completely different.

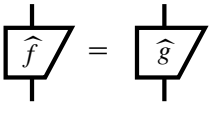
*6.6.1 Doubling General Process Theories**

There is nothing particularly special about **linear maps** that enabled us to construct a doubled process theory. Given any process theory **p** that admits string diagrams, we obtain a new process theory $\mathcal{D}(\mathbf{p})$, again admitting string diagrams, which is made up of all processes of the form:

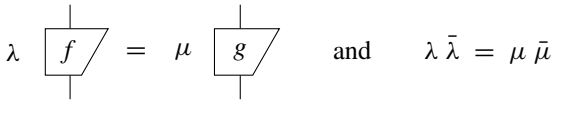


where f can be any process in **p**. Many of the theorems we proved in this chapter and later also don't depend on linear maps or have natural generalisations. For example, as we already noted in Remark 6.19, the proof that doubling eliminates 'generalised global phases' was purely diagrammatic, so it immediately generalises.

Theorem 6.102 Let f and g be processes in any process theory that admits string diagrams, then we have:



if and only if there exist λ and μ such that:

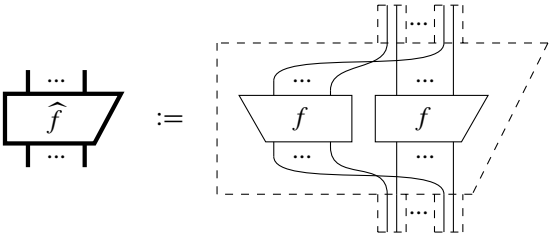


Furthermore, any process theory of the form $\mathcal{D}(\mathbf{p})$ admits a discarding map, so things such as causality, purification, and Stinespring dilation all still make perfect sense.

Thus many of the concepts we develop apply not just to quantum theory, but can be used to study a whole family of 'quantum-like' theories. The upshot of doing so is that, first, by drawing contrasts with the other theories, this provides insights into the true identity of quantum theory. Second, as discussed in Section 1.2.4, the future theories of physics might not be based on **linear maps** but on some new kind of beast.

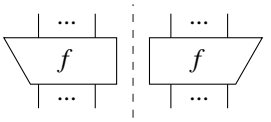
Exercise* 6.103 Characterise $\mathcal{D}$(**relations**). What are its states? What are its general processes? What does causality mean?

There are in fact many equivalent ways to build such a doubled process theory. Notably the 'twisted' version we use in this book:

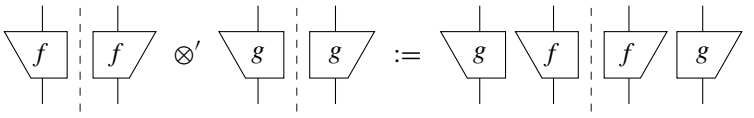


is easy to work with diagrammatically because it uses the same parallel composition as the non-doubled theory and keeps the 'two halves' of each system close together. This comes in particularly handy in Chapter 8 when we start modelling classical-quantum interaction.

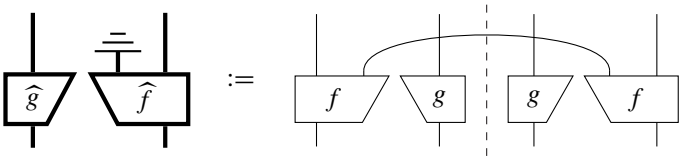
The original version of this construction (which is usually called the *CPM-construction*) avoids this twisting:



but at the price of needing to define a new parallel composition in the doubled category:



Now, the 'two halves' of each system might be far apart, but you could imagine folding the paper along the dotted line to see which systems should be combined together to make thick wires. A discarding process then introduces a connection between the left half and the right half:



This version of doubling is in fact easier to define using category theory (where it originates), but its diagrams start to become messy fairly quickly.

*6.6.2 Axiomatizing Doubling**

Can we tell if a process theory is the result of doubling? You might say: 'That's easy, just check if the wires are thick!' Okay, so maybe we should rephrase the question. What we

are looking for here is an axiomatic characterisation of process theories of the form $\mathcal{D}(\mathbf{p})$ for some $\mathbf{p}$.

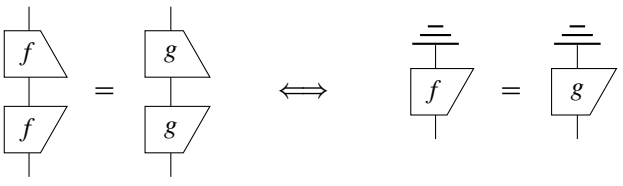
The crucial aspect of that theory is that it has discarding processes that can be used to purify any process. Note that, since we don't know (in advance) that a process theory $\mathbf{P}$ is of the form $\mathcal{D}(\mathbf{p})$, we needed to give a characterisation of 'pure' processes that doesn't refer to doubling. Here is the result.

Definition 6.104 A *discarding structure* for a process theory admitting string diagrams $\mathbf{P}$ consists of:

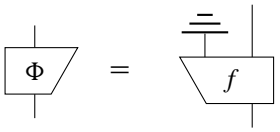
- a subtheory $\mathbf{p} \subseteq \mathbf{P}$ of 'pure processes' that includes all of the types from $\mathbf{P}$, and
- for each type A, a discarding effect:

which are subject to the following axioms:

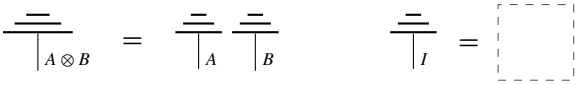
1. All processes f, g in $\mathbf{p}$ satisfy:



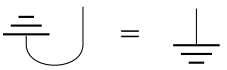
2. Every process Φ in $\mathbf{P}$ can be *purified*; that is, there always exists a pure process f in $\mathbf{p}$ such that Φ arises by discarding part of its output:



3. Discarding two systems means discarding each of them; discarding nothing means doing nothing:



and the transpose of discarding is the same as its adjoint:

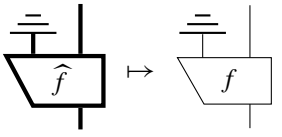


Each process theory with a discarding structure is equivalent to one that arises from doubling some process theory.

Theorem 6.105 If **P** has a discarding structure, then:

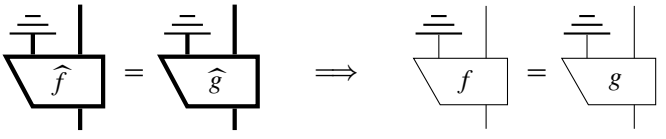
$$\mathcal{D}(\mathbf{p}) \cong \mathbf{P}$$

Proof We prove this by showing a bijection between the processes in $\mathcal{D}(\mathbf{p})$ and the processes in **P**. Since **p** has the same types as **P**, there is an obvious correspondence between the types $\widehat{A}$ in $\mathcal{D}(\mathbf{p})$ and A in **P**. Similarly, since **p** is a subtheory of **P**, there is an obvious correspondence between pure processes $\widehat{f}$ in $\mathcal{D}(\mathbf{p})$ and their counterparts f in **P**. But what about the impure processes? For these, we first purify, then interpret (the normal) discarding process as the discarding structure from **P**:

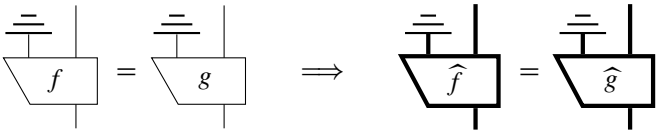


(6.82)

For this map to be well defined, it should not depend on the particular choice of purification; i.e. it should be the case that:



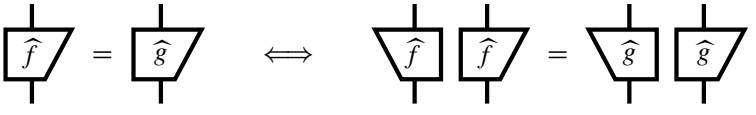
and for it to be injective, the reverse implication should hold:



Both of these can be proven using Axiom 1 in Definition 6.104 and a healthy dose of process–state duality. Axiom 2 guarantees that the mapping (6.82) is surjective. From these facts and Axiom 3, it is then fairly straightforward to show that this mapping also preserves string diagrams. We leave it to the reader to fill in the details. □

One would expect the converse to hold as well; namely, process theories of the form $\mathcal{D}(\mathbf{p})$ should carry a discarding structure, where the pure processes are precisely those of the form $\widehat{f}$ for f in **p**. It turns out this is almost true, provided that the pure processes in $\mathcal{D}(\mathbf{p})$ satisfy one extra property.

Exercise 6.106 Show that if for pure processes in $\mathcal{D}(\mathbf{p})$ we have that:



then $\mathcal{D}(\mathbf{p})$ carries a discarding structure. This condition can be read as: 'the doubled theory has no global phases'.

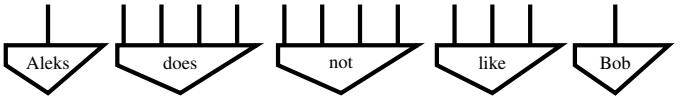
*6.6.3 And Now for Something Completely Different**

We now know that string diagrams and the doubling construction are useful tools for expressing quantum processes. However, surprisingly, they have been employed in a totally different area: computational linguistics!

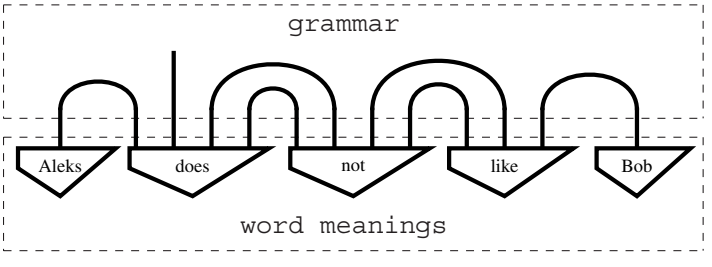
A typical problem in computational linguistics is to take a pair of words and compute how similar they are in meaning. For example, 'dog' and 'hound' should be very similar, 'dog' and 'cat' less similar (but at least they are still both animals), and 'dog' and 'taxes' should not be similar at all. The most common way to solve this problem is to compute vectors for the words 'dog', 'hound', and 'taxes', often by scanning a big body of text and 'learning' the vectors automatically using some techniques from artificial intelligence. Then, as we already saw back in Section 4.3.3, we can measure the similarity (i.e. 'commonality' or 'overlap') just by taking the inner product.

This method is now standard, and you will find it under the hood of basically any software that deals with human language somehow (web search, translation, targeted advertising, etc.). But where things really start to get interesting is in trying to extend this from words to sentences.

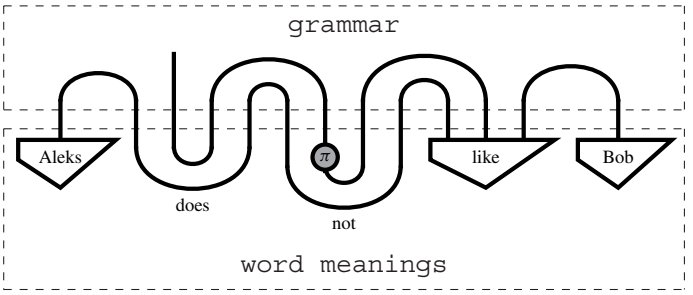
Consider this possible outcome of co-authoring a textbook:



Each word is a state, and different kinds of words (nouns, transitive verbs, intransitive verbs, etc.) have different types (hence the different numbers of wires). Then, to compute the meaning of this sentence, all we need to do is wire these words together. For this, we need a set of rules for how to combine the words together, but that's just called *grammar*:



We treat some words as 'black boxes', whereas others may be represented as diagrams themselves, indicating how they interact with other words:



at which point we can start to do the types of diagrammatic calculations we've seen all throughout this book.

So what does doubling everything do for us? Well, just as in the quantum case, it makes room for 'impurity'. Does this have a place in this model of language? Of course it has! Many words are ambiguous; for example, Bob may refer to a folk singer, a reggae singer, a cartoon builder, a cartoon sponge, or one of the authors of this book. We can use mixing to represent this ambiguity.

6.7 Historical Notes and References

Max Planck (1900) is considered as the originator of quantum theory. He made his discovery when studying how to get the most light out of light bulbs with the minimal amount of electricity, a job he was doing for the electricity companies. For this purpose, he studied the radiation of light by a perfect absorber of light and realised that the energy carried by light could only be emitted in certain packages, called *quanta*. These packages of energy are given by *Planck's law*, $E = h\nu$, where h is *Planck's constant*, ν is the frequency of the light, and E is the corresponding energy package.

Then, Heisenberg (1925) formulated a predecessor of quantum theory using indexed sets of 'transition amplitudes'; Born and Jordan (1925) realised that Heisenberg was actually talking about (infinite) complex matrices; Schrödinger (1926) instead was speaking in waves; Dirac (1926) realised that these were actually one and the same thing; and finally von Neumann (1932) put everything together in the language of Hilbert space. The Born rule was introduced by Born (1926). Many more were involved in all of this at an earlier stage, and many obtained Nobel Prizes for their respective contributions, including Einstein, Bohr, Compton, de Broglie, Fermi, and Pauli. The Bloch sphere was introduced by Felix Bloch (1946) as a representation of the states of two-dimensional quantum systems, and spin-$\frac{1}{2}$ systems in particular.

Von Neumann (1927b) introduced density matrices, which were used to represent mixed states. Quantum maps, to which one usually refers as completely positive maps, were first proposed as the general dynamics of quantum systems in Sudarshan et al. (1961). As mathematical entities these maps had been around quite a bit longer as maps between C*-algebras (see e.g. Paulsen, 2002). It is in this context that Stinespring dilation first appeared (Stinespring, 1955). Kraus decomposition can be found in Kraus (1983) and the no-broadcasting theorem first appeared in Barnum et al. (1996). Another 'generalised

no-broadcasting theorem' is Barnum et al. (2007), which, rather than process theories, concerns generalised probabilistic theories.

As we shall discuss in Section 7.3.2, von Neumann's formulation of quantum theory was quite different from the presentation of quantum theory as a process theory. The doubling construction of Section 6.1 that turns linear maps into pure quantum maps was introduced in Coecke (2007), including the proof of Theorem 6.17. The generalisation to all quantum maps from Section 6.2 was introduced by Selinger (2007) as the *CPM-construction*. The idea that this can be done by adding the discarding process was put forward by Coecke (2008), and the 'ground' notation for discarding was introduced in Coecke and Perdrix (2010). The axiomatisation of the CPM-construction of Section 6.6.2 also first appeared in Coecke (2008).

Quantum processes without classical input are also called *quantum instruments*, a notion originated by Davies (1976) and further developed by Ozawa (1984).

The causality postulate that governs the passage from quantum maps to quantum processes of Section 6.4 was only recently identified as a core principle of quantum theory, by Chiribella et al. (2010, 2011), where it was one of a series of axioms from which quantum theory was reconstructed. In fact, it was defined in terms of probabilities adding up to one, from which uniqueness of causal effects was derived. There is, however, a discrepancy between the terminology in those papers and in ours, in that those papers use the term 'deterministic' to mean 'causal' in the sense of equation (6.64), so this may also refer to what we call non-deterministic quantum processes.

The characterisation of all teleportation protocols as in Section 6.4.6 appeared first in Werner (2001). The first proof of the non-signalling theorem for quantum theory can be found in Ghirardi et al. (1980). The derivation of non-signalling from the causality principle as in Section 6.3.2 is taken from Coecke (2014b). A similar result is also in Fritz (2014) and Henson et al. (2014). The derivation of covariance from causality as in Section 6.3.3 is taken from Coecke and Lal (2013), in which earlier covariance results of Markopoulou (2000) and Blute et al. (2003) were generalised to causal process theories.

Abramsky and Coecke (2004) started the axiomatic approach that takes composition of systems as its starting point. There were earlier approaches that tried to take composition of systems as a starting point of a new axiomatic approach (see e.g. Coecke, 2000), but with no real success. Two recent reconstructions of quantum theory that take diagrams as their backbone are Chiribella et al. (2011) and Hardy (2012). We already provided references for the earlier axiomatic approaches to quantum theory at the end of Chapter 1. The alternative definition of purity hinted at in Remark 6.81 is taken from Chiribella (2014). In fact, Chiribella only defined purity of states in this way. It turns out that extending this definition to general processes only makes sense for quantum systems, and not classical ones. Since classical systems can be broadcast, not even the bare wire is 'pure' using this alternative definition, even though 'doing nothing' is about as 'pure' as it gets!

The use of quantum maps for describing meaning in natural language was initiated in Coecke et al. (2010c); a direct comparison with quantum teleportation can be found in Clark et al. (2014); and the interpretation of ambiguity as mixing is taken from Piedeleu et al. (2015).