Lecture 8: A process theory of linear maps

Summary.
(1) Basis and sums in string diagrams.
(2) Matrices.
(3) Linear maps as processes, Hilbert spaces as their types.

Luís Soares Barbosa,
Univ. Minho *(Informatics Dep.)* & Inl *(Quantum Software Engineering Group)*

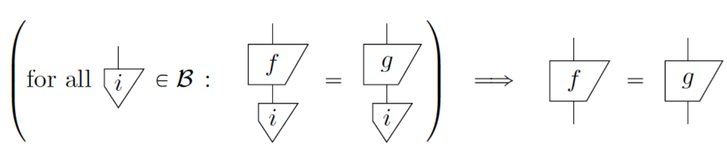
Introduction.

This lecture is devoted to the construction of a particular process theory — that of *linear maps*, which plays an important role in the study of quantum phenomena. Each type in this theory comes with a finite orthonormal basis. A basis determines the type *dimension*. The existence of at least a type of every dimension, and the fact that all processes of a given type can be 'summed' are other characteristics of this process theory. What these notions — *basis*, *dimension* and *sum* mean in the framework of string diagrams needs to be clarified. Finally, scalars in the theory of linear maps are complex numbers, thus providing an example of a theory whose scalars are not self-conjugate. Types in this theory are, as expected, Hilbert spaces.

Although a main goal of the approach introduced in this series of lectures is to replace matrices of complex numbers by string diagrams when reasoning about quantum computation, the whole theory of linear processes can indeed be rephrased in terms of matrices. The basic observation is that to show two processes equal in this theory it is enough to look at scalars. Matrices, on their own, provide another example of a process theory, which is equivalent, as a category, to the process theory of linear maps (recall the notion of equivalent categories from a previous lecture)[1].

Basis.

In the process theory of relations states of a type A can be regarded as subsets of A. However, two relations can be shown equal by analysing only their composition with a particular class of states — those corresponding to singleton sets. In general, a basis for a type A in a process theory is a minimal set of states $\mathcal{B}$ such that, for all processes f and g,

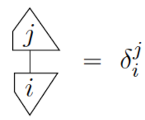


(1)

[1]Pictures are taken from Coecke and Kissinger book, *Picturing Quantum processes*, CUP, 2017.

The *dimension* of a type A is the size of its smallest basis.

A basis is *orthonormal* if its elements are perfectly distinguishable by testing, i.e. they do not overlap. Formally, if for all i and j,

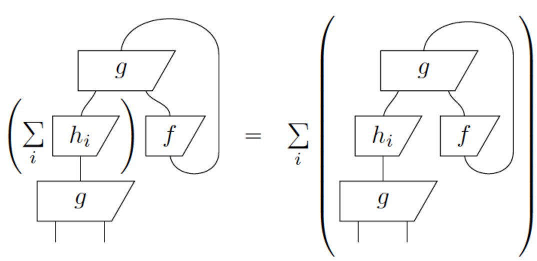


Exercise 1

Show that if an orthonormal set of states of a given type satisfies condition (1) for all pairs of processes, then it is the minimal set to do it, and therefore forms an orthonormal basis (ONB).

Sum.

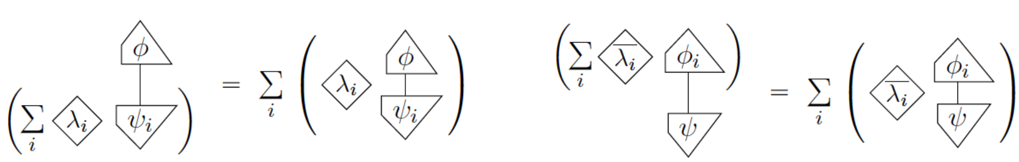
A process theory has *sums* if its processes are equipped with an Abelian monoid structure $(+, \mathsf{0})$ which preserves adjoints and distributes over diagrams. The last condition subsumes distributivity of sums with respect to parallel and sequential composition. Graphically, it means that summations can always be pulled outside as in the following example:



Actually, the summation symbol can freely move around the diagram, just like a scalar.

Exercise 2

Which well-known properties are expressed by the following equations?



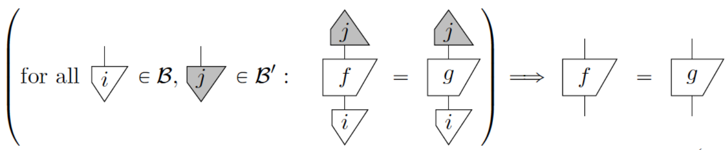
$\boxed{\textbf{Exercise } 3}$

What is sum in the process theory of relations? Show that scalars in this process theory come with a 'sum' and a 'multiplication' related by a distributive law, the basic structure one associates to a reasonable notion of a number.

From processes to matrices and back.

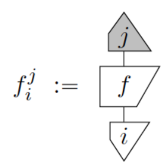
$\boxed{\textbf{Exercise } 4}$

Suppose one wants to prove two processes $f, g : A \longrightarrow B$ to be equal. Assuming bases $\mathcal{B}$, $\mathcal{B}'$ for A and B, respectively, show that

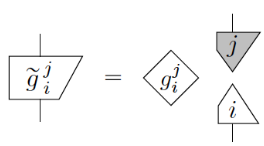


The result above is interesting: equality of processes is established by comparing a set of scalars. If both bases are orthonormal, the scalars that uniquely identify a process are the entries of a matrix representing it,

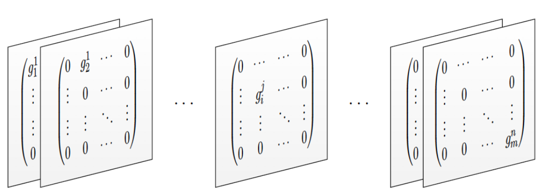
$$\begin{pmatrix} f_1^1 & f_2^1 & \cdots & f_m^1 \\ f_1^2 & f_2^2 & \cdots & f_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ f_1^n & f_2^n & \cdots & f_m^n \end{pmatrix}$$



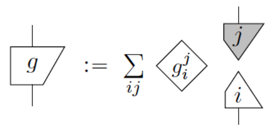
Dually, processes can be recovered from arbitrary matrices. Suppose a matrix whose entries are g_i^j. Choose orthonormal bases and compute, for each matrix entry, a process



whose matrix has exactly g_i^j in position (i, j). The original matrix corresponds to the sum of all those processes

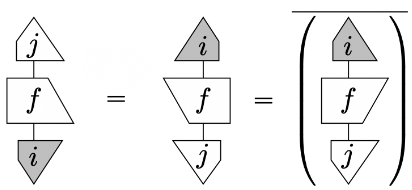


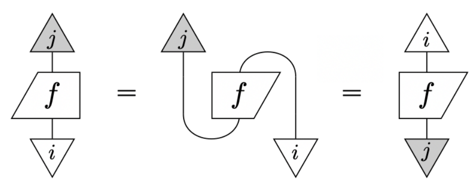
Thus,

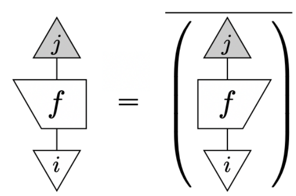


Let us denote by $\mathbf{f}$ the matrix representation of process f. Then, the adjoint, transpose and conjugate of a process have straightforward matrix representations, as follows,

$$(\mathbf{f}^\dagger)^{\mathsf{j}}_{\mathsf{i}} := \overline{(\mathbf{f}^{\mathsf{i}}_{\mathsf{j}})} \qquad (\mathbf{f}^{\mathsf{T}})^{\mathsf{j}}_{\mathsf{i}} := \mathbf{f}^{\mathsf{i}}_{\mathsf{j}}; \qquad \overline{\mathbf{f}}^{\mathsf{j}}_{\mathsf{i}} := \overline{(\mathbf{f}^{\mathsf{j}}_{\mathsf{i}})}$$







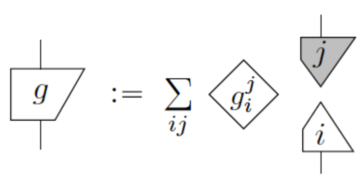
Notice that the last two cases resort to a fundamental assumption that bases are self-conjugate. Without it the resulting matrices would be expressed in terms of the conjugates of the original bases.

Exercise 5

Show that the matrix representing the sum of a collection of processes $\Sigma_{\mathsf{k}}\mathsf{f}_{\mathsf{k}}$ is the sum $\sum_{\mathsf{k}} \mathbf{f_k}$ of matrices $\mathbf{f_k}$, where

$$(\Sigma_{\mathsf{k}}\mathbf{f}_{\mathsf{k}})^{\mathsf{j}}_{\mathsf{i}} = \sum_{\mathsf{k}} (\mathbf{f}_{\mathsf{k}})^{\mathsf{j}}_{\mathsf{i}}$$

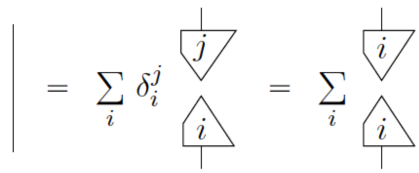
Given suitable orthonormal bases for its input and output, a process written as



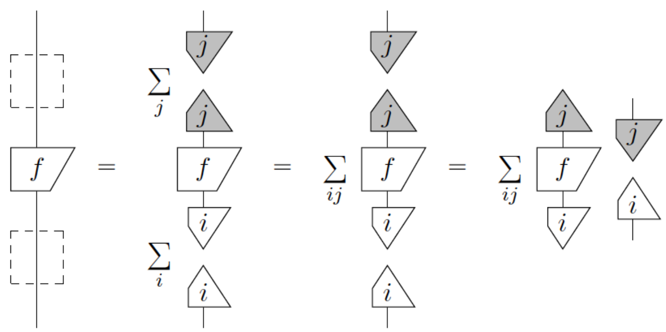
is said to be in *matrix form*. Clearly, the corresponding matrix is

$$\begin{pmatrix} g_1^1 & g_2^1 & \cdots & g_m^1 \\ g_1^2 & g_2^2 & \cdots & g_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ g_1^n & g_2^n & \cdots & g_m^n \end{pmatrix}$$

The entry (i, j) in the matrix representing the identity process is $\delta_{i,}j$. The whole matrix boils down to the identity matrix. Translating to matrix form we recover the well-known *resolution of the identity* decomposition:



The matrix form for the identity process gives a handy way to compute the matrix form for an arbitrary process as follows

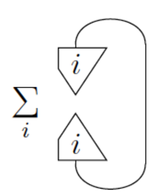


Exercise 6

Prove that a set off states form an orthonormal basis iff, for all i,

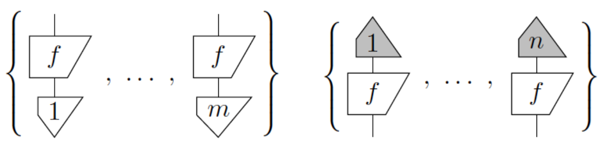
Exercise 7

Compute

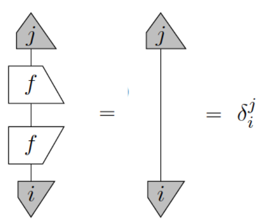


and interpret the result in the process theory of relations and of linear maps.

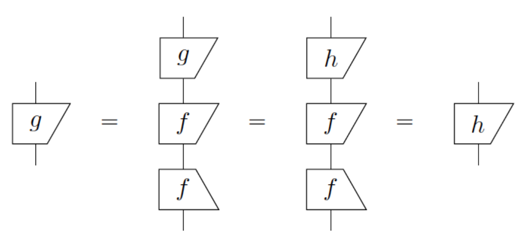
We need now a smooth way to recognise some specific processes in terms of their matrices. Thus, first note that, given orthonormal bases for the input and output types of a process f, the columns, respectively rows, of **f** are the matrices corresponding to the following sets of states, respectively of effects,



In this framework, a unitary process f sends orthonormal bases to orthonormal bases. Indeed, for any basis $\mathcal{B}$, the set of states $\mathsf{f} \cdot \mathsf{i}$ is orthonormal because, being unitary, f is an isometry which entails



It remains to show that these states form a basis. This amounts to show that two processes g and h are equal when they agree on all of these states. Assume $\mathsf{g} \cdot \mathsf{f} \cdot \mathsf{i} = \mathsf{h} \cdot \mathsf{f} \cdot \mathsf{i}$ for all i. Since the set of states i form a basis, we have $\mathsf{g} \cdot \mathsf{f} = \mathsf{h} \cdot \mathsf{f}$. Finally, f being unitary yields



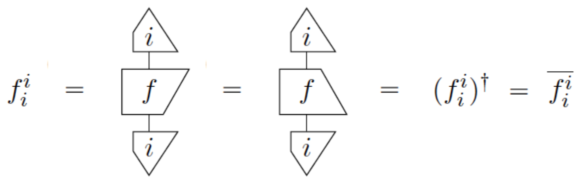
Exercise 8

Show that both columns and rows of a matrix corresponding to a unitary process form orthonormal bases.

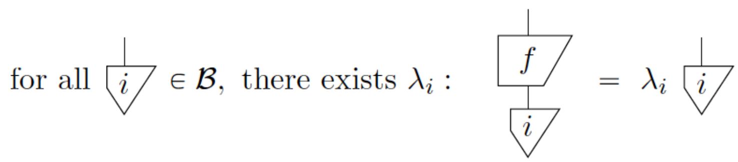
The matrices corresponding the self-adjoint processes are easily recognised

$$\begin{pmatrix} f_1^1 & f_2^1 & \cdots & f_n^1 \\ \overline{f_2^1} & f_2^2 & \cdots & f_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \overline{f_n^1} & \overline{f_n^2} & \cdots & f_n^n \end{pmatrix}$$

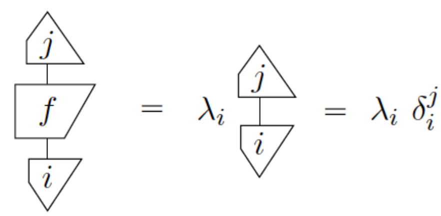
where elements of the diagonal are self-conjugate:



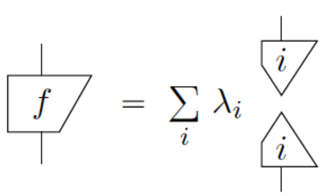
The same is not true for e.g. matrices corresponding to positive processes. However, for the class of so-called *diagonalisable* processes, self-adjointness, positivity, and projectors can be identified from the scalars on the diagonal of the corresponding matrix. A process f is *diagonalisable* if there exists a orthonormal basis $\mathcal{B}$ such that all basis states are eigenstates of f, i.e.



This entails



Thus its matrix is a diagonal matrix:



$$\leftrightarrow \quad \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Exercise 9

Show that f is self-adjoint iff all elements in the diagonal are self-conjugate; positive iff all positive; and a projector if they are all positive and satisfy $\lambda_i = (\lambda_i)^2$.

Exercise 10

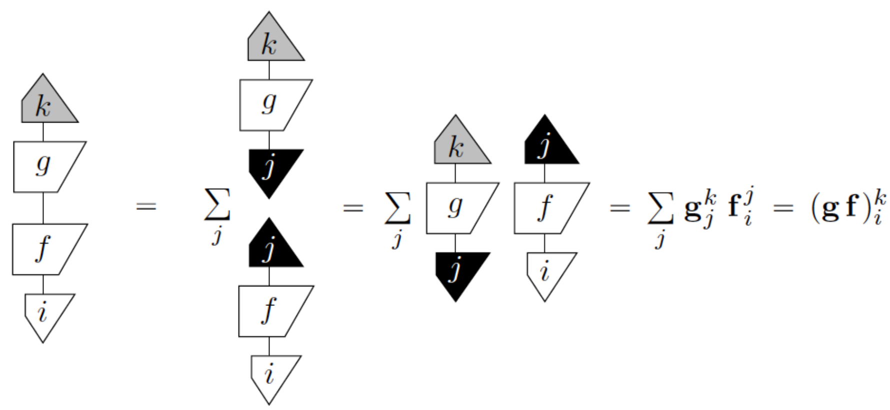
What is the matrix form of the trace of a process? And of a partial trace?

Matrices for circuits.

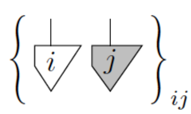
Defining sequential and parallel composition of matrices is in order so that we can have an idea of how circuits translate into composition of matrices. Thus, the matrix representation of $\mathsf{g} \cdot \mathsf{f}$ is the matrix product $\mathbf{gf}$, i.e.

$$(\mathbf{g} \cdot \mathbf{f})_{\mathsf{i}}^{\mathsf{k}} := \Sigma_{\mathsf{j}} \mathbf{g}_{\mathsf{j}}^{\mathsf{k}} \mathbf{f}_{\mathsf{i}}^{\mathsf{j}}$$

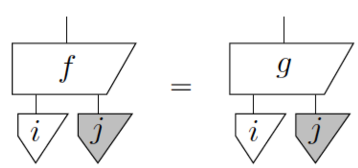
Actually,



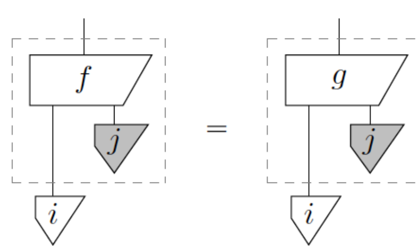
To compute the matrix for $\mathsf{f} \otimes \mathsf{g}$ we first need a way to build a basis for a tensor type $\mathsf{A} \otimes \mathsf{B}$, from the bases for $\mathcal{B}$, $\mathcal{B}'$, of A and B. The set



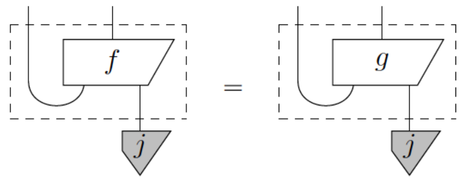
does the job. To show it forms indeed a basis, assume that any pair of processes with input type $\mathsf{A} \otimes \mathsf{B}$ agrees on all the states above, i.e.



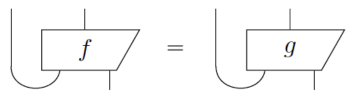
or equivalently



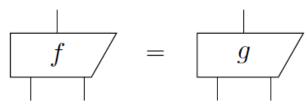
Since $\mathcal{B}$ is a basis,



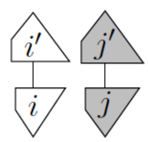
yielding, because $\mathcal{B}'$ is a basis as well,



and, therefore,



Orthonormality, i.e.



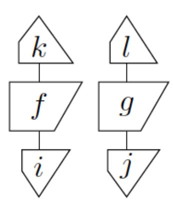
$$= \delta_i^{i'}\delta_j^{j'}$$

comes from the original bases being orthonormal as well.

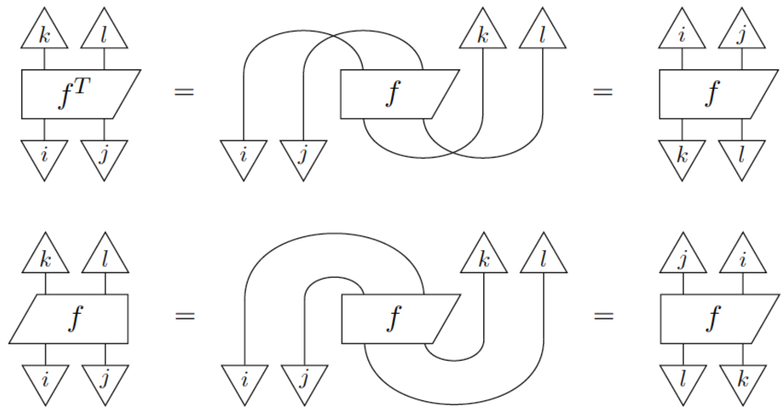
The matrix representation of $\mathsf{f}\otimes\mathsf{g}$ is the Kronecker product of $\mathbf{g}$ and $\mathbf{f}$, i.e.

$$(\mathbf{f}\otimes\mathbf{g})^{\mathsf{kl}}_{\mathsf{ij}} := \mathbf{f}^{\mathsf{k}}_{\mathsf{i}}\mathbf{g}^{\mathsf{j}}_{\mathsf{l}}$$

the entries being given by



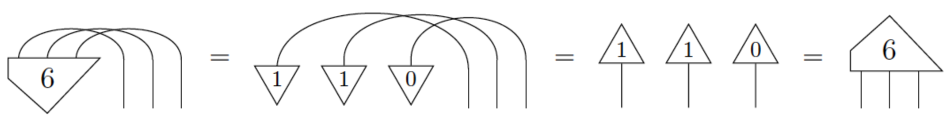
Finally, a note on transposition. Recall the two notion of transposition discussed in a previous lecture:



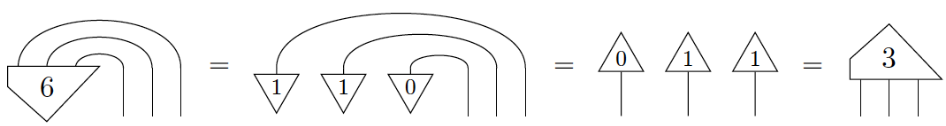
Algebraic transposition (and thus algebraic conjugate) is the one capturing the notion of transposition in linear algebra. The more usual notion of a rotation, reverses the order, besides interchanging, as expected, superscripts and subscripts. Actually, transposition of matrices yields

$$\mathbf{f}_{\mathsf{ij}}^{\mathsf{kl}} \rightsquigarrow \mathbf{f}_{\mathsf{kl}}^{\mathsf{ij}} \quad \text{rather than} \quad \mathbf{f}_{\mathsf{ij}}^{\mathsf{kl}} \rightsquigarrow \mathbf{f}_{\mathsf{lk}}^{\mathsf{ji}}$$

Moreover, the algebraic transpose also keeps the most significant basis state on the left, and the least significant on the right, as illustrated by



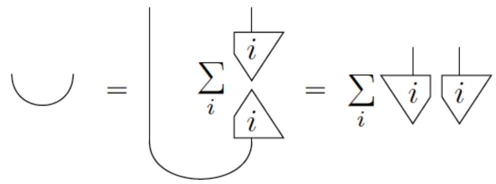
versus



Matrices for string diagrams.

Going from circuits to diagrams entails the need to define matrices for cups and caps. For any orthonormal basis, we obtain from the matrix form of the identity (the so-called resolution of the identity)

because



The dual case comes from taking the adjoint. Note that this result applies to any orthonormal basis, thus, in particular, to the conjugate of the original basis, yielding an equivalent characterisation

Using both (equivalent) definitions the yanking laws follows immediately. In practice, however, these details can be avoided by focusing only on self-conjugate basis; graphically,

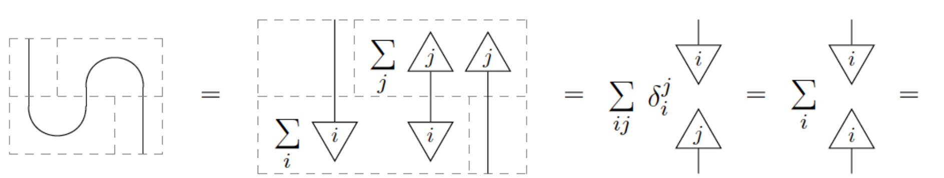
Once re-written in the usual Dirac notation these equations emerge as the usual definition of the Bell state and effect.

$\boxed{\textbf{Exercise } 11}$

Which are the matrices for caps and cup for systems with dimension 2?

$\boxed{\textbf{Exercise } 12}$

Actually one could take this as the definition of cups and caps for a type A. The yanking laws are obviously satisfied; for example,

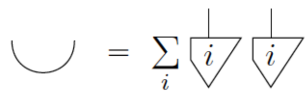


Complete the verification for the remaining laws.

One may always choose, in a unique way, caps and cups to make a given orthonormal basis self-conjugate.

Exercise 13

Show that an orthonormal basis is self-conjugate iff



Matrices as a paradigmatic process theory.

Having discussed how matrices can be combined into string diagrams, one may easily conclude that matrices not only provide a convenient way to represents processes (as e.g. relations or linear maps), but give rise to process theories themselves. Indeed, given a suitable notion of scalar N with the usual 'arithmetic' structure and a possibly trivial conjugation, matrices over N form a process theory. Types in such a theory are just natural numbers standing for the number of columns and rows — the role of types in a process theory is to mediate sequential composition which, in this case, boils down to matrix multiplication.

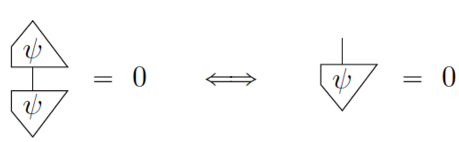
Moreover, every process theory with scalars N, in which each type has a finite orthonormal basis, processes of the same type admit sums, and there is at least one type of each dimension, is equivalent to the process theory of matrices over N. The word equivalent stands for an equivalence of categories as introduced earlier in this course.

Scalars for linear maps.

In the introduction to this lecture, we point out the three new elements to consider to develop a process theory of linear maps: a notion of basis, a sum operation and a specific choice of scalars as the complex numbers. Let us complete the picture exploring this last element. Complex numbers provide scalars in which conjugation is not trivial. This means that the adjoint of a linear map does not coincide with the transpose.

Exercise 14

Show that the positive-definiteness condition in the definition of the inner product (cf, Lecture 5), i.e.



requires the adjoint to be defined as the conjugate transpose.

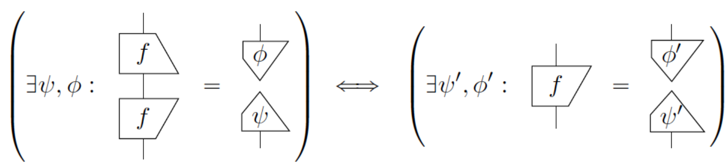
Thus, this is a process theory in which the three basic notions in string diagrams — transpose, adjoint, conjugate — are indeed distinct.

This choice of scalars becomes really handy. An example is the well-known *dimension theorem* which states that in any process theory which admits a matrix representations (as discussed above) and whose scalars form a field[2], all bases for a given type are the same size.

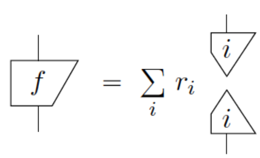
Another important example is the following result which states that the composition $f^\dagger \cdot f$ provides a tool to decide whether process f is $\cdot$-separable (just as positive-definiteness mentioned above expresses the ability of number $\phi^\dagger \cdot \phi$ to detect if state ϕ is 0). Formally,

Exercise 15

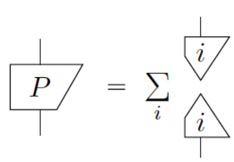
Show that for any linear map f,



This property is a consequence of a fundamental result: the *spectral theorem* which reads as follows: All self-adjoint linear maps are diagonalisable, i.e. there is an orthonormal basis such that



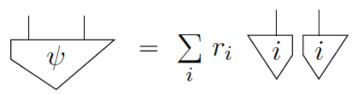
where all r_i are real numbers. Moreover, if f is positive (resp. a projector) all $r_i \geq 0$ (belong to $\{0, 1\}$). This is precisely the theorem that, for an orthonormal set, allows to express each projector P as



[2]A a field is a set on which addition, subtraction, multiplication, and division are defined and behave as the corresponding operations on real or complex numbers do.

Exercise 16

As a corollary of the spectral theorem show that for all self-conjugate bipartite states in the theory of linear maps there exists an orthonormal basis such that



Similarly, all r_i are greater or equal to zero for $\otimes$-positive states. A crucial observation in this proof is that, under the process-state duality, self-conjugate (resp, $\otimes$-positive) bipartite states correspond to self-adjoint (resp. positive) processes, as discussed in Lecture 5.

The property discussed in Exercise 15 does not hold for relations.

Exercise 17

Provide a counter-example.

The reason is the same why the spectral theorem fails in the theory of relations: there are not enough scalars in the theory to guarantee a diagonal representation for every relation.

The process theory of linear maps.

Summing up, this theory consists of all processes expressed by string diagrams such that

- There is at least a type for each dimension $n \in \mathbb{N}$ and an orthonormal basis for each type;
- Processes of the same type admit sums;
- Scalars are the complex numbers $\mathbb{C}$.

A Hilbert space is simply a type in this theory, denoted by $\mathbb{C}^n$ for a n-dimensional type. Notice that the familiar, set-theoretic definition of an Hilbert space can easily be recovered. The emphasis here, however, as in Category Theory in general, is on stressing the relevance of processes (arrows) with respect to types (objects).

The real relevant result is the following: linear maps are complete for string diagrams. Formally, an equation between string diagrams holds for all Hilbert spaces and linear maps if and only if the string diagrams are the same. Clearly this does not happen for relations: the equation

$$\lambda\,\lambda \;=\; \lambda$$

holds in the process theory of relations although the string diagrams corresponding to both sides are different. Recall that string diagrams are equal if they can be deformed into each other.

Dealing with sums and bases categorically.

The introduction of sums in a process theory is categorically captured by the notion of *enriched category*, which denotes a category whose hom-sets (i.e. sets of arrows between a pair of objects) have some additional structure. In the case of interest here, the appropriate structure is that a commutative monoid. This means that $\mathsf{Hom}(A, B)$, for every pair of objects A and B, forms a commutative monoid and the monoid structure is compatible with sequential composition, i.e.

$$\begin{aligned} 0 \cdot f &= f \cdot 0 = 0 \\ h \cdot (f + g) &= (h \cdot f) + (h \cdot g) \\ (f + h) \cdot f &= (g \cdot f) + (h \cdot f) \end{aligned}$$

At the object level, one may define a notion somehow corresponding to sums of arrows as follows: A category enriched over a commutative monoid has *biproducts* if, for each pair of objects A_1 and A_2, there exists an object $A \oplus B$ and a pair of maps

$$\begin{aligned} \pi_i &: A_1 \oplus A_2 \longrightarrow A_i \\ \iota_i &: A_i \longrightarrow A_1 \oplus A_2 \end{aligned}$$

such that

$$\iota_1 \cdot \pi_1 + \iota_2 \cdot \pi_2 = \mathrm{id}_{A_1 \oplus A_2}$$

and

$$\pi_i \cdot \iota_j = \begin{cases} \mathrm{id}_{A_j} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

In a dagger compact closed category, as introduced in Lecture 5, enriched over a commutative monoid, biproducts are defined in terms of maps

$$\iota_i : A_i \longrightarrow A_1 \oplus A_2$$

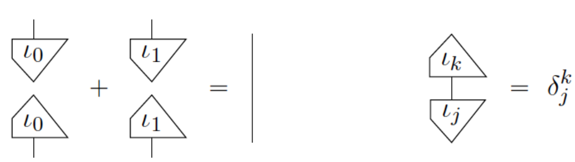
satisfying

$$\iota_1 \cdot (\iota_1)^\dagger + \iota_2 \cdot (\iota_1)^\dagger = \mathrm{id}_{A_1 \oplus A_2}$$

and

$$(\iota_1)_i^\dagger \cdot \iota_j = \begin{cases} \mathrm{id}_{A_j} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Interestigly enough, taking $A_1 = A_2 = \mathrm{I}$ these conditions boil down to



which means that the set of states labelled by ι_j forms a two-dimensional orthonormal basis.

Being an associative operator, one may define biproducts of an arbitrary, finite, number of objects. In particular, objects that can be written as

$$A \;\cong\; I \oplus I \oplus I \cdots \oplus I$$

have a n-dimensional orthonormal basis.

An object that cannot be written as a biproduct of two non-zero objects is classified as *irreducible*. The existence of irreducible objects are therefore connected to the existence of orthonormal bases. The reason linear maps (and relations) have orthonormal bases for every type is that in their theories the only irreducible is the trivial object I.