Categories and Quantum Informatics: Hilbert spaces

Chris Heunen

Spring 2018

We introduce our main example category **Hilb** by recalling in some detail the mathematical formalism that underlies quantum theory: (complex) vector spaces, inner products, orthonormal bases, linear maps, matrices, dimensions, and dual spaces. We then introduce the adjoint of a linear map between Hilbert spaces, and define the terms unitary, isometry, partial isometry, and positive. We also define the tensor product of Hilbert spaces, and introduce the Kronecker product of matrices.

0.9 Vector spaces

A vector space is a collection of elements that can be added to one another, and scaled.

Definition 0.1 (Vector space)**.** A *vector space*, is a set V with a chosen element $0 \in V$, an addition operation $+\colon V \times V \longrightarrow V$, and a scalar multiplication operation $\cdot\colon \mathbb{C} \times V \longrightarrow V$, satisfying the following properties for all $u, v, w \in V$ and $a, b \in \mathbb{C}$:

- *additive associativity*: $u + (v + w) = (u + v) + w$;
- *additive commutativity*: $u + v = v + u$;
- *additive unit*: $v + 0 = v$;
- *additive inverses*: there exists a $-v \in V$ such that $v + (-v) = 0$;
- *additive distributivity*: $a \cdot (u + v) = (a \cdot u) + (a \cdot v)$
- *scalar unit*: $1 \cdot v = v$;
- *scalar distributivity*: $(a + b) \cdot v = (a \cdot v) + (b \cdot v)$;
- *scalar compatibility*: $a \cdot (b \cdot v) = (ab) \cdot v$.

The prototypical example of a vector space is $\mathbb{C}^n$, the cartesian product of n copies of the complex numbers.

Definition 0.2 (Linear map, antilinear map)**.** A *linear map* is a function $f\colon V \longrightarrow W$ between vector spaces, with the following properties, for all $u, v \in V$ and $a \in \mathbb{C}$:

$$f(u + v) = f(u) + f(v) \tag{16}$$
$$f(a \cdot v) = a \cdot f(v) \tag{17}$$

An *anti-linear map* is a function that also satisfies (16), but instead of (17), has the additional property

$$f(a \cdot v) = a^* \cdot f(v), \tag{18}$$

where the star denotes complex conjugation.

We can use these definitions to build a category of vector spaces.

Definition 0.3 (Vect, FVect). The category **Vect** of vector spaces and linear maps is defined as follows:

- **objects** are complex vector spaces;
- **morphisms** are linear functions;
- **composition** is composition of functions;
- **identity morphisms** are identity functions.

We define the category **FVect** to be the restriction of **Vect** to those vector spaces that are isomorphic to $\mathbb{C}^n$ for some natural number n; these are also called *finite-dimensional*, see Definition 0.5 below.

A *kernel* of a morphism $A \xrightarrow{f} B$ in a category is an equaliser of f and the *zero morphism* $A \xrightarrow{0} B$. Any morphism $f\colon V \to W$ in **Vect** has a kernel, namely the inclusion of $\ker(f) = \{v \in V \mid f(v) = 0\}$ into V. Hence kernels in the categorical sense coincide precisely with kernels in the sense of linear algebra.

Definition 0.4. The *direct sum* of vector spaces V and W is the vector space $V \oplus W$, whose elements are pairs (v, w) of elements $v \in V$ and $w \in W$, with entrywise addition and scalar multiplication.

Direct sums are both products and coproducts in the category **Vect**.

0.10 Bases and matrices

One of the most important structures we can have on a vector space is a *basis*. They give rise to a the notion of dimension of a vector space, and let us represent linear maps using matrices.

Definition 0.5. For a vector space V, a family of elements $\{e_i\}$ is *linearly independent* when every element $a \in V$ can be expressed as a finite linear combination $a = \sum_i a_i e_i$ with *coefficients* $a_i \in \mathbb{C}$ in at most one way. It is a *basis* if additionally any $a \in V$ can be expressed as such a finite linear combination.

Every vector space admits a basis, and any two bases for the same vector space have the same cardinality. This is not clear, but quite nontrivial to show.

Definition 0.6. The *dimension* of a vector space V, written $\dim(V)$, is the cardinality of any basis. A vector space is *finite-dimensional* when it has a finite basis.

If vector spaces V and W have bases $\{d_i\}$ and $\{e_j\}$, and we fix some order on the bases, we can represent a linear map $V \xrightarrow{f} W$ as the matrix with $\dim(W)$ rows and $\dim(V)$ columns, whose entry at row i and column j is the coefficient $f(v_j)_i$. Composition of linear maps then corresponds to matrix multiplication (7). This directly leads to a category.

Definition 0.7. The skeletal category $\mathbf{Mat}_{\mathbb{C}}$ is defined as follows:

- **objects** are natural numbers $0, 1, 2, \ldots$;
- **morphisms** $n \to m$ are matrices of complex numbers with m rows and n columns;
- **composition** is given by matrix multiplication;
- **identities** $n \xrightarrow{\mathrm{id}_n} n$ are given by n-by-n matrices with entries 1 on the main diagonal, and 0 elsewhere.

This theory of matrices is 'just as good' as the theory of finite-dimensional vector spaces. This can be made precise using the category theory we have developed.

Proposition 0.8. *There is an equivalence of categories* $\mathbf{Mat}_{\mathbb{C}} \to \mathbf{FVect}$, *sending* $n \mapsto \mathbb{C}^n$, *and a matrix to its associated linear map.*

Proof. Because every finite-dimensional complex vector space H is isomorphic to $\mathbb{C}^{\dim(H)}$, the functor R is essentially surjective on objects. It is full and faithful since there is an exact correspondence between matrices and linear maps for finite-dimensional vector spaces. $\square$

For a square matrix, the trace is an important operation.

Definition 0.9. For a square matrix with entries m_{ij}, its *trace* is the number $\sum_i m_{ii}$ given by the sum of the entries on the main diagonal.

0.11 Hilbert spaces

Hilbert spaces are structures that are built on vector spaces. The extra structure lets us define angles and distances between vectors, and is used in quantum theory to calculate probabilities of measurement outcomes.

Definition 0.10. An *inner product* on a complex vector space V is a function $\langle - | - \rangle \colon V \times V \to \mathbb{C}$ that is:

- *conjugate-symmetric*: for all $v, w \in V$,

$$\langle v|w\rangle = \langle w|v\rangle^*, \tag{19}$$

- *linear* in the second argument: for all $u, v, w \in V$ and $a \in \mathbb{C}$,

$$\langle v|a \cdot w\rangle = a \cdot \langle v|w\rangle, \tag{20}$$
$$\langle u|v + w\rangle = \langle u|v\rangle + \langle u|w\rangle; \tag{21}$$

- *positive definite*: for all $v \in V$,

$$\langle v|v\rangle \geq 0, \tag{22}$$
$$\langle v|v\rangle = 0 \Rightarrow v = 0. \tag{23}$$

Definition 0.11. For a vector space with inner product, the *norm* of an element v is $\|v\| := \sqrt{\langle v|v\rangle}$, a nonnegative real number.

The complex numbers carry a canonical inner-product structure given by

$$\langle a|b\rangle := a^* b, \tag{24}$$

where $a^* \in \mathbb{C}$ denotes the complex conjugate of $a \in \mathbb{C}$.

This norm satisfies the triangle inequality $\|v+w\| \leq \|v\|+\|w\|$ by virtue of the Cauchy-Schwarz inequality $|\langle v|w\rangle|^2 \leq \langle v|v\rangle \cdot \langle w|w\rangle$, that holds in any vector space with an inner product. Thanks to these properties, it makes sense to think of $\|u - v\|$ as the distance between vectors u and v.

A Hilbert space is an inner product space in which it makes sense to add infinitely many vectors in certain cases.

Definition 0.12. A *Hilbert space* is a vector space H with an inner product that is *complete* in the following sense: if a sequence $v_1, v_2, \ldots$ of vectors satisfies $\sum_{i=1}^{\infty} \|v_i\| < \infty$, then there is a vector v such that $\|v - \sum_{i=1}^{n} v_i\|$ tends to zero.

Every finite-dimensional vector space with inner product is necessarily complete. Any vector space with an inner product can be completed to a Hilbert space by adding in appropriate limit vectors.

There is a notion of bounded map between Hilbert spaces that makes use of the inner product structure. The idea is that for each map there is some maximum amount by which the norm of a vector can increase.

Definition 0.13 (Bounded linear map)**.** A linear map $f\colon H \longrightarrow K$ between Hilbert spaces is *bounded* when there exists a number $b \in \mathbb{R}$ such that $\|f(v)\| \leq b \cdot \|v\|$ for all $v \in H$.

Every linear map between finite-dimensional Hilbert spaces is bounded.

Hilbert spaces and bounded linear maps form a category. For the purposes of modelling phenomena in quantum theory, this category will be the main example that we use throughout the book.

Definition 0.14 (**Hilb**, **FHilb**)**.** The category **Hilb** of Hilbert spaces and bounded linear maps is defined as follows:

- **objects** are Hilbert spaces;
- **morphisms** are bounded linear maps;
- **composition** is composition of linear maps as ordinary functions;
- **identity morphisms** are given by the identity linear maps.

We define the category **FHilb** to be the restriction of **Hilb** to finite-dimensional Hilbert spaces.

This definition is perhaps surprising, especially in finite dimensions: since every linear map between Hilbert spaces is bounded, **FHilb** is an equivalent category to **FVect**. In particular, the inner products play no essential role. We will see later how inner products can be modelled categorically, using the idea of *daggers*.

Hilbert spaces have a more discerning notion of basis.

Definition 0.15 (Basis, orthogonal basis, orthonormal basis)**.** For a Hilbert space H, an *orthogonal basis* is a family of elements $\{e_i\}$ with the following properties:

- they are *pairwise orthogonal*, *i.e.* $\langle e_i | e_j \rangle = 0$ for all $i \neq j$;
- every element $a \in H$ can be written as an infinite linear combination of e_i; *i.e.* there are *coefficients* $a_i \in \mathbb{C}$ for which $\|a - \sum_{i=1}^{n} a_i e_i\|$ tends to zero.

It is *orthonormal* when additionally $\langle e_i | e_i \rangle = 1$ for all i.

Any orthogonal family of elements is automatically linearly independent. For finite-dimensional Hilbert spaces, the ordinary notion of basis as a vector space is still useful, as given by Definition 0.5. Hence once we fix (ordered) bases on finite-dimensional Hilbert spaces, linear maps between them correspond to matrices, just as with vector spaces. For infinite-dimensional Hilbert spaces, however, having a basis for the underlying vector space is rarely mathematically useful.

If two vector spaces carry inner products, we can give an inner product to their direct sum, leading to the direct sum of Hilbert spaces.

Definition 0.16. The *direct sum* of Hilbert spaces H and K is the vector space $H \oplus K$, made into a Hilbert space by the inner product $\langle (v_1, w_1) | (v_2, w_2) \rangle = \langle v_1 | v_2 \rangle + \langle w_1 | w_2 \rangle$.

Direct sums provide both products and coproducts for the category **Hilb**. Hilbert spaces have the good property that any closed subspace can be complemented. That is, if the inclusion $U \hookrightarrow V$ is a morphism of **Hilb** satisfying $\|u\|_U = \|u\|_H$, then there exists another inclusion morphism $U^\perp \hookrightarrow V$ of **Hilb** with $V = U \oplus U^\perp$. Explicitly, $U^\perp$ is the *orthogonal subspace* $\{v \in V \mid \forall u \in U \colon \langle u | v \rangle = 0\}$.

0.12 Adjoints

The inner product gives rise to the *adjoint* of a bounded linear map.

Definition 0.17. For a bounded linear map $f\colon H \longrightarrow K$, its *adjoint* $f^\dagger\colon K \longrightarrow H$ is the unique linear map with the following property, for all $u \in H$ and $v \in K$:

$$\langle f(u)|v\rangle = \langle u|f^\dagger(v)\rangle. \tag{25}$$

The existence of the adjoint follows from the Riesz representation theorem for Hilbert spaces, which we do not cover here. It follows immediately from (25) by uniqueness of adjoints that they also satisfy the following properties:

$$(f^\dagger)^\dagger = f, \tag{26}$$

$$(g \circ f)^\dagger = f^\dagger \circ g^\dagger, \tag{27}$$

$$\mathrm{id}_H{}^\dagger = \mathrm{id}_H. \tag{28}$$

Taking adjoints is an *anti-linear* operation.

We can use adjoints to define various specialized classes of linear maps.

Definition 0.18. A bounded linear map $H \xrightarrow{f} K$ between Hilbert spaces is:

- *self-adjoint* when $f = f^\dagger$;
- a *projection* when $f = f^\dagger$ and $f \circ f = f$;
- *unitary* when both $f^\dagger \circ f = \mathrm{id}_H$ and $f \circ f^\dagger = \mathrm{id}_K$;
- an *isometry* when $f^\dagger \circ f = \mathrm{id}_H$;
- a *partial isometry* when $f^\dagger \circ f$ is a projection;
- and *positive* when $f = g^\dagger \circ g$ for some bounded linear map $H \xrightarrow{g} K$.

The following notation is standard in the physics literature.

Definition 0.19 (Bra, ket)**.** Given an element $v \in H$ of a Hilbert space, its *ket* $\mathbb{C} \xrightarrow{|v\rangle} H$ is the linear map $a \mapsto av$. Its *bra* $H \xrightarrow{\langle v|} \mathbb{C}$ is the linear map $w \mapsto \langle v|w\rangle$.

You can check that $|v\rangle^\dagger = \langle v|$. The reason for this notation is demonstrated by the following calculation:

$$\left(\mathbb{C} \xrightarrow{|v\rangle} H \xrightarrow{\langle w|} \mathbb{C}\right) = \left(\mathbb{C} \xrightarrow{\langle w|\circ|v\rangle} \mathbb{C}\right) = \left(\mathbb{C} \xrightarrow{\langle w|v\rangle} \mathbb{C}\right) \tag{29}$$

In the final expression here, we identify the number $\langle w|v\rangle$ with the linear map that sends $1 \mapsto \langle w|v\rangle$. We see that the inner product (or 'bra-ket') $\langle w|v\rangle$ breaks into a composite of a bra $\langle w|$ and a ket $|v\rangle$. Originally due to Paul Dirac, this is traditionally called *Dirac notation*.

The correspondence between $|v\rangle$ and $\langle v|$ leads to the notion of a dual space.

Definition 0.20. For a Hilbert space H, its *dual Hilbert space* H^* is the vector space $\mathbf{Hilb}(H, \mathbb{C})$.

A Hilbert space is isomorphic to its dual in an anti-linear way: the map $H \longrightarrow H^*$ given by $|v\rangle \mapsto \varphi_v = \langle v|$ is an invertible anti-linear function. The inner product on H^* is given by $\langle \varphi_v|\varphi_w\rangle_{H^*} = \langle v|w\rangle_H$, and makes the function $|v\rangle \mapsto \langle v|$ bounded.

For some bounded linear maps, we can define a notion of trace.

Definition 0.21 (Trace, trace class)**.** When it converges, the *trace* of a positive linear map $f\colon H \longrightarrow H$ is given by $\mathrm{Tr}(f) := \sum \langle e_i|f(e_i)\rangle$ for any orthonormal basis $\{e_i\}$, in which case the map is called *trace class*.

If the sum converges for one orthonormal bases, then with some effort one can prove that it converges for all orthonormal bases, and that the trace is independent of the chosen basis. Also, in the finite-dimensional case, the trace defined in this way agrees with the matrix trace of Definition 0.9.

0.13 Tensor products

The tensor product is a way to make a new vector space out of two given ones. With some work the tensor product can be constructed explicitly, but it is only important for us that it exists, and is defined up to isomorphism by a universal property. If U, V and W are vector spaces, a function $f\colon U \times V \to W$ is called *bilinear* when it is linear in each variable: when the function $u \mapsto f(u, v)$ is linear for each $v \in V$, and the function $v \mapsto f(u, v)$ is linear for each $u \in U$.

Definition 0.22. The *tensor product of vector spaces* U and V is a vector space $U \otimes V$ together with a bilinear function $f\colon U \times V \to U \otimes V$ such that for every bilinear function $g\colon U \times V \to W$ there exists a unique linear function $h\colon U \otimes V \to W$ such that $g = h \circ f$.

$$\begin{tikzcd} U \times V \arrow[r, "\text{(bilinear) } f"] \arrow[dr, "\text{(bilinear) } g"'] & U \otimes V \arrow[d, dashed, "h \text{ (linear)}"] \\ & W \end{tikzcd}$$

The function f usually stays anonymous and is written as $(u, v) \mapsto u \otimes v$. It follows that arbitrary elements of $U \otimes V$ take the form $\sum_{i=1}^{n} a_i u_i \otimes v_i$ for $a_i \in \mathbb{C}$, $u_i \in U$, and $v_i \in V$. The tensor product also extends to linear maps. If $f_1\colon U_1 \to V_1$ and $f_2\colon U_2 \to V_2$ are linear maps, there is a unique linear map $f_1 \otimes f_2\colon U_1 \otimes U_2 \to V_1 \otimes V_2$ that satisfies $(f_1 \otimes f_2)(u_1 \otimes u_2) = f_1(u_1) \otimes f_2(u_2)$ for $u_1 \in U_1$ and $u_2 \in U_2$. In this way, the tensor product becomes a functor $\otimes\colon \mathbf{Vect} \times \mathbf{Vect} \to \mathbf{Vect}$.

Definition 0.23. The *tensor product of Hilbert spaces* H and K is the following Hilbert space $H \otimes K$: take the tensor product of vector spaces; give it the inner product $\langle u_1 \otimes v_1 | u_2 \otimes v_2 \rangle = \langle u_1 | u_2 \rangle_H \cdot \langle v_1 | v_2 \rangle_K$; complete it. If $H \xrightarrow{f} H'$ and $K \xrightarrow{g} K'$ are bounded linear maps, then so is the continuous extension of the tensor product of linear maps to a function that we again call $f \otimes g\colon H \otimes K \to H' \otimes K'$. This gives a functor $\otimes\colon \mathbf{Hilb} \times \mathbf{Hilb} \to \mathbf{Hilb}$.

If $\{e_i\}$ is an orthonormal basis for Hilbert space H, and $\{f_j\}$ is an orthonormal basis for K, then $\{e_i \otimes f_j\}$ is an orthonormal basis for $H \otimes K$. So when H and K are finite-dimensional, there is no difference between their tensor products as vector spaces and as Hilbert spaces.

Definition 0.24 (Kronecker product)**.** When finite-dimensional Hilbert spaces H_1, H_2, K_1, K_2 are equipped with fixed ordered orthonormal bases, linear maps $H_1 \xrightarrow{f} K_1$ and $H_2 \xrightarrow{g} K_2$ can be written as matrices. Their tensor product $H_1 \otimes H_2 \xrightarrow{f \otimes g} K_1 \otimes K_2$ corresponds to the following block matrix, called their *Kronecker product*:

$$(f \otimes g) := \begin{pmatrix} \big(f_{11} g\big) & \big(f_{12} g\big) & \cdots & \big(f_{1n} g\big) \\ \big(f_{21} g\big) & \big(f_{22} g\big) & \ldots & \big(f_{2n} g\big) \\ \vdots & \vdots & \ddots & \vdots \\ \big(f_{m1} g\big) & \big(f_{m2} g\big) & \ldots & \big(f_{mn} g\big) \end{pmatrix}. \tag{30}$$