Chapter 1

Hilbert spaces and operators

We begin by studying finite dimensional Hilbert spaces. We do so abstractly rather than just the concrete Hilbert space $\mathbb{C}^n$, as many Hilbert spaces will arise in higher linear algebra which do not exactly look like $\mathbb{C}^n$. We study the abstract complex $*$-algebra $B(H)$ of linear operators on a Hilbert space H rather than $M_n(\mathbb{C})$ for the same reason. Some proofs may certainly be simplified by replacing $B(H)$ with $M_n(\mathbb{C})$ and a 'unitary algebra' (finite dimensional $\mathrm{C}^*/\mathrm{W}^*$-algebra) with a multimatrix algebra. However, the proofs here are really operator algebraic in nature and many can be adapted to the infinite dimensional setting (with more work).

1.1 Finite dimensional vector spaces

In this section, we rapidly recall the basic notions from finite dimensional linear algebra that will be used in this book. Many proofs are omitted, as this material is assumed prerequisite. We always work over the field of complex numbers.

We assume the reader is already familiar with complex vector spaces, together with the notions of what it means for a set of vectors to be *linearly independent* or to *span* a vector space V. A set of vectors which spans and is linearly independent is called a *basis* for V.

Exercise 1.1.1. Prove that if $B \subset V$ is a basis, then every element of V can be written uniquely as a (finite) linear combination of elements of B.

Exercise 1.1.2. Show that every linear map $V \to W$ is completely determined by where it sends a basis of V. Deduce that one may define a linear map simply by specifying where it sends a basis.

The first important theorem in linear algebra concerns when subsets of $\mathbb{C}^n$ can span or be linearly independent. It is proven using the *Gaussian elimination algorithm*, which we omit from this book.

Theorem 1.1.3 — Suppose $S = \{v_1, \ldots, v_k\} \subset \mathbb{C}^n$.

- If S spans $\mathbb{C}^n$, then $k \geq n$.
- If S is linearly independent, then $k \leq n$.

Hence every basis of $\mathbb{C}^n$ has exactly n elements.

Proof. Consider the $n \times k$ matrix A whose columns are the v_i

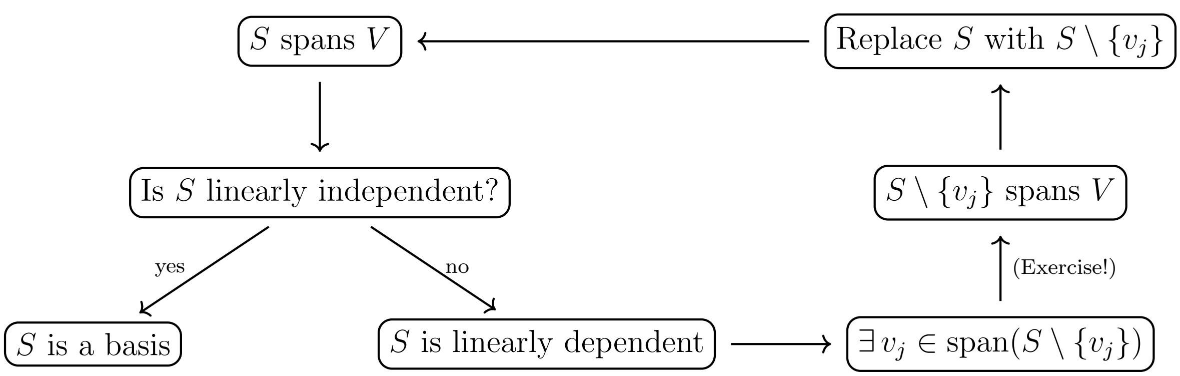
$$A := \left[v_1 \middle| \cdots \middle| v_k \right], \tag{1.1.4}$$

and consider the matrix $\mathrm{ref}(A)$ in *row echelon form* obtain by applying the Gaussian Elimination Algorithm to A. The v_i span $\mathbb{C}^n$ if and only if there is a leading one/pivot in every row of $\mathrm{ref}(A)$, which implies $k \geq n$. The v_i are linearly independent if and only if there is a leading one/pivot in every column of $\mathrm{ref}(A)$, which implies $k \leq n$. $\square$

In this book, we will be primarily concerned with *finite dimensional* vector spaces. We begin by considering those spaces which are *finitely generated*, i.e., V admits a finite spanning set S.

Corollary 1.1.5 (Contraction) — Given a finite spanning set S for V, there is a basis $B \subset S$ for V. Thus every finitely generated vector space admits a finite basis.

Proof. The proof consists of the following algorithm, presented as a flowchart.



Since S is finite, the algorithm terminates. $\square$

Definition 1.1.6 (Coordinate map) — Suppose $B = \{e_1, \ldots, e_n\}$ is a basis for V. The *coordinate* isomorphism $[\cdot] = [\cdot]_B : V \to \mathbb{C}^n$ is given by

$$V \ni v = \sum \lambda_i e_i \longmapsto [v]_B := \begin{bmatrix} \lambda_1 \\ \ddots \\ \lambda_n \end{bmatrix} \in \mathbb{C}^n. \tag{1.1.7}$$

Theorem 1.1.8 — Suppose $S, T \subset V$ such that S spans V and T is linearly independent. Then $|T| \leq |S|$.

Proof. Let $B \subset S$ be a basis, which exists by Corollary 1.1.5. Enumerate $T = \{v_1, \ldots, v_k\}$. The map

$$\mathbb{C}^k \ni \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{bmatrix} \longmapsto \sum \lambda_i v_i \longmapsto \left[\sum \lambda_i v_i\right]_B \in \mathbb{C}^n$$

is a composite of two injective maps and is thus injective. Using Theorem 1.1.3, we conclude that $|T| = k \leq n = |B| \leq |S|$. $\square$

Definition 1.1.9 (Dimension) — By Theorem 1.1.8, every basis of a finitely generated vector space V has the same cardinality. This cardinality is called the *dimension* of V, denoted $\dim(V)$.

Exercise 1.1.10. Dualize the flowchart in Corollary 1.1.5 above to prove that every linearly independent subset of a finite dimensional vector space can be extended to a basis.

Definition 1.1.11 (Coordinates for linear maps) — Given bases $B = \{e_j\}_{j=1}^n$ for V and $C = \{f_k\}_{k=1}^m$ for W and a linear map $x : V \to W$, we can define a matrix $[x] = [x]_B^C \in M_{m \times n}(\mathbb{C})$ by

$$[x]_B^C = \left[[xv_1]_C \middle| \cdots \middle| [xv_n]_C \right]$$

That is, the j-th column of $[x]_B^C$ is given by $[x]_j := [xv_j]_C$. Here, the columns correspond to the basis of the source, and the rows correspond to the basis of the target. Under this isomorphism, composition of linear operators corresponds to matrix multiplication.

That is, the following diagram commutes.

$$\begin{tikzcd}
(V,\{e_j\}_{j=1}^n) \arrow[r, "x"] \arrow[d, "{[\,\cdot\,]}"'] \arrow[rr, bend left, "yx"] & (W,\{f_k\}_{k=1}^m) \arrow[r, "y"] \arrow[d, "{[\,\cdot\,]}"] & (X,\{g_\ell\}_{\ell=1}^p) \arrow[d, "{[\,\cdot\,]}"] \\
\mathbb{C}^n \arrow[r, "{[x]}"] \arrow[rr, bend right, "{[y][x]}"'] & \mathbb{C}^m \arrow[r, "{[y]}"] & \mathbb{C}^p
\end{tikzcd}$$

Thus studying maps between vectors spaces and operations between them is studying matrices and their operations. We may thus replace a vector space V with $\mathbb{C}^n$ and the space of linear maps $\mathrm{Hom}(V \to W)$ with $M_{m\times n}(\mathbb{C})$ without loss of generality.

Definition 1.1.12 — Given a vector space V, its *dual space* is $V^\vee := \mathrm{Hom}(V \to \mathbb{C})$, the space of linear maps $V \to \mathbb{C}$. There is a canonical map from V to the *double dual* $V^{\vee\vee}$ given by $v \mapsto [\mathrm{eval}_v : f \mapsto f(x)]$. When V is finite dimensional, this map is a linear isomorphism, i.e., it is invertible.

Construction 1.1.13 — Given a basis $\{v_1,\ldots,v_n\}$ for V, the *dual basis* $\{v_1^\vee,\ldots,v_n^\vee\}$ for $V^\vee$ is defined by the formula

$$v_i^\vee(v_j) = \delta_{i=j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

One verifies the dual basis is actually a basis.

Warning 1.1.14 — Even though $\dim(V) = \dim(V^\vee)$ so that $V \cong V^\vee$, there is no canonical isomorphism. That is, one gets a linear isomorphism for any choice of basis for V, but choosing different bases gives fundamentally different isomorphisms. Indeed, the space of ordered bases for V is $GL(V)$, which is not contractible. (It is harder to describe the space of unordered bases, so we will not do so here.)

Remark 1.1.15. One can define the coordinate map $[\,\cdot\,]_B : V \to \mathbb{C}^n$ in terms of the dual basis for $B = \{v_1,\ldots,v_n\}$ as $[v]_j := v_j^\vee(v)$.

Definition 1.1.16 — There are two fundamental subspaces associated to a linear map $x : V \to W$:

- $\ker(x) = \{v \in V \mid xv = 0\} \subseteq V$, and

- $\mathrm{im}(x) = xV = \{xv \mid v \in V\} \subseteq W$.

The *rank* of x is $\mathrm{rank}(x) := \dim(\mathrm{im}(x))$, and the *nullity* of x is $\mathrm{nullity}(x) := \dim(\ker(x))$.

Theorem 1.1.17 (Rank-Nullity) — For a linear operator $x : V \to W$ where V, W are finite dimensional,

$$\mathrm{rank}(x) + \mathrm{nullity}(x) = \dim(V).$$

Proof. By using coordinates, without loss of generality, $V = \mathbb{C}^n$ and $W = \mathbb{C}^m$. Using the Gaussian Elimination Algorithm, $\mathrm{rank}(x)$ is the number of columns of $\mathrm{ref}([x])$ with leading ones/pivots, and $\mathrm{nullity}(x)$ is the number of columns of $\mathrm{ref}([x])$ witout leading ones/pivots. Clearly these sum to $n = \dim(\mathbb{C}^n)$. $\square$

1.2 Normed spaces

Before delving into our study of Hilbert spaces, we begin with the notion of a normed space, which should be viewed as a vector space equipped with a notion of *length* for vectors.

Definition 1.2.1 — A *normed space* is a vector space V equipped with a *norm*, i.e., a function $\|\cdot\| : V \to [0, \infty)$ satisfying:

- (subadditive) $\|\eta + \xi\| \leq \|\eta\| + \|\xi\|$ for all $\eta, \xi \in V$,
- (homogeneous) $\|\lambda\eta\| = |\lambda|\|\eta\|$ for all $\eta \in V$ and $\lambda \in \mathbb{C}$,
- (definite) $\|\eta\| = 0$ if and only if $\eta = 0$.

The norm, which should be viewed as the a length function, induces a notion of *distance* between vectors.

Definition 1.2.2 — A *metric* on a vector space V consists of a function $d \colon V \times V \to [0, \infty)$ such that:

- (triangle inequality) $d(\xi_1, \xi_2) \leq d(\xi_1, \eta) + d(\eta, \xi_2)$ for all $\xi_1, \eta, \xi_2 \in V$,
- (symmetry) $d(\eta, \xi) = d(\xi, \eta)$ for all $\eta, \xi \in V$,
- (definite) $d(\eta, \xi) = 0$ if and only if $\eta = \xi$.

Exercise 1.2.3. Prove that every norm $\|\cdot\|$ on V gives rise to a metric $d(\eta, \xi) := \|\eta - \xi\|$. Conversely, show that a metric d on V arises in this way if and only if it satisfies:

- (homogeneous) $d(\lambda\eta, \lambda\xi) = |\lambda| d(\eta, \xi)$ for all $\eta, \xi \in H$ and $\lambda \in \mathbb{C}$.

Hint: Set $\|\eta\| := d(0, \eta)$ for $\eta \in H$.

Exercise 1.2.4. Prove that $\|\xi\|_1 := \sum |\xi_i|$ and $\|\xi\|_\infty := \max |\xi_i|$ are norms on $\mathbb{C}^n$. Then prove that

$$\|\xi\|_\infty \leq \|\xi\|_1 \leq n \cdot \|\xi\|_\infty \qquad \forall \xi \in \mathbb{C}^n.$$

Note: The norm $\|\cdot\|_1$ is endearingly known as the taxi-cab norm. *The corresponding metric $d(x, y) = \|x - y\|_1$ is sometimes called the* Manhattan metric.

Definition 1.2.5 — Suppose V, W are vector spaces. A *(linear) map* $x \colon V \to W$ is a function which is

- (linear) $x(\lambda\eta + \xi) = \lambda x\eta + x\xi$ for all $\eta, \xi \in H$ and $\lambda \in \mathbb{C}$.

When $W = \mathbb{C}$, we say that x is a *linear functional*.

It is straightforward to verify that a linear operator is invertible (an inverse exists) if and only if it bijective. Observe that $x : V \to W$ is bijective if and only if $\ker(x) = 0$ and $\operatorname{im}(x) = W$, where

$$\ker(x) := \{\eta \in V \mid x\eta = 0\} \qquad \text{and} \qquad \operatorname{im}(x) := xV = \{x\eta \mid \eta \in H\}.$$

Example 1.2.6 — Every matrix $A \in M_{m\times n}(\mathbb{C})$ gives a linear map $L_A : \mathbb{C}^n \to \mathbb{C}^m$ by left multiplication.

Definition 1.2.7 — Suppose V, W are normed spaces. For a linear map $x : V \to W$, we define its *operator norm* by

$$\|x\|_{\text{op}} := \sup\{\|A\xi\| \mid \xi \in V \text{ with } \|\xi\| = 1\}.$$

Exercise 1.2.8. Prove that $\|\cdot\|_{\text{op}}$ is a norm on the space of linear maps from $V \to W$ when V, W are normed spaces such that $\|x\xi\| \leq \|x\|_{\text{op}} \cdot \|\xi\|$ for all $\xi \in V$.

Exercise 1.2.9. Suppose V, W are normed spaces and $x : V \to W$ is linear. Prove that if $C > 0$ such that $\|x\xi\| \leq C \cdot \|\xi\|$ for all $\xi \in V$, then $\|x\|_{\text{op}} \leq C$. Deduce that $\|x\|_{\text{op}}$ is the infemum of such $C > 0$.

Lemma 1.2.10 — The following are equivalent for a sequence $(\eta_k) \subset \mathbb{C}^n$ and $\eta \in \mathbb{C}^n$.

(1) $\|\eta_k - \eta\| \to 0$ for some norm $\|\cdot\|$ on $\mathbb{C}^n$, and

(2) $\eta_k \to \eta$ component-wise.

Hence all norms on $\mathbb{C}^n$ induce the same *topology*.

Proof. Translating (η_k) by η, we may assume $\eta = 0$.

(1)⇒(2): Given $\xi \in \mathbb{C}^n = M_{1\times n}(\mathbb{C})$, its i-th component is given by $E_i\xi$ where $E_i \in M_{1\times n}(\mathbb{C})$ is the matrix with a 1 in its i-th component and zero elsewhere. By Exercise 1.2.8, left multiplication by E_i is a continuous map from $(\mathbb{C}^n, \|\cdot\|) \to (\mathbb{C}, |\cdot|)$.

(2)⇒(1): Observe $\|\eta_k\|_\infty = \max |E_i\eta_k| \to 0$. □

Proposition 1.2.11 — All norms on $\mathbb{C}^n$ are equivalent. That is, if $\|\cdot\|_1, \|\cdot\|_2$ are two norms on $\mathbb{C}^n$, there is a $C > 0$ such that $C^{-1}\|\cdot\|_2 \leq \|\cdot\|_1 \leq C\|\cdot\|_2$.

Proof. Without loss of generality, we may assume that $\|\cdot\|_2$ is our favorite norm on $\mathbb{C}^n$. We fix our favorite for which we know that the unit ball is compact. (Ours is $\|\cdot\|_\infty$, for which the unit ball is $[-1,1]^n$.) Then the unit sphere (the $x \in \mathbb{C}^n$ such that $\|x\|_2 = 1$) is also compact. By Lemma 1.2.10, the $\|\cdot\|_1$-image of the $\|\cdot\|_2$-unit sphere (which does not contain the zero vector) is a compact, connected subset of $(0,\infty)$, which is some closed and bounded subinterval. Pick $C > 0$ such that both

$$C^{-1} \leq \min_{\|x\|_2=1} \|x\|_1 \qquad \text{and} \qquad \max_{\|x\|_2=1} \|x\|_1 \leq C.$$

Then whenever $x \in \mathbb{C}^n$ is non-zero,

$$C^{-1} \leq \left\| \frac{x}{\|x\|_2} \right\|_1 \leq C \qquad \Longleftrightarrow \qquad C^{-1}\|x\|_2 \leq \|x\|_1 \leq C\|x\|_2.$$ □

1.3 Hilbert spaces

Definition 1.3.1 — A *Hilbert space* is a vector space H equipped with a positive definite *inner product*, i.e., a function $\langle\cdot|\cdot\rangle : H \times H \to \mathbb{C}$ such that

- (linear in second variable) $\langle\eta|\lambda\xi_1 + \xi_2\rangle = \lambda\langle\eta|\xi_1\rangle + \langle\eta|\xi_2\rangle$ for all $\eta, \xi_1, \xi_2 \in H$ and $\lambda \in \mathbb{C}$,
- (anti-symmetric) $\overline{\langle\eta|\xi\rangle} = \langle\xi|\eta\rangle$ for all $\eta, \xi \in H$,
- (positive definite) $\langle\eta|\eta\rangle \geq 0$ for all $\eta \in H$ with equality if and only if $\eta = 0$.

(Since H was assumed to be finite dimensional, there is no completeness condition!) The *length* of $\eta \in H$ is $\|\eta\| := \sqrt{\langle\eta|\eta\rangle}$. (The length is a norm; this is Exercise 1.3.7 below.)

Example 1.3.2 — The space $\mathbb{C}^n$ is a Hilbert space with $\langle\eta|\xi\rangle = \sum_{j=1}^n \overline{\eta_j}\xi_j$. Under the identification of $\mathbb{C}^n = M_{n\times 1}(\mathbb{C})$, $\langle\eta|\xi\rangle = \eta^\dagger\xi$, where $\dagger$ denotes taking the conjugate transpose.

Exercise 1.3.3. A *sesquilinear form* on a complex vector space V is a function $(\cdot|\cdot) : V^2 \to \mathbb{C}$ which is linear in the second variable and anti-linear in the first variable, i.e.,

$$(\lambda\eta_1 + \eta_2|\xi) = \overline{\lambda}(\eta_1|\xi) + (\eta_2|\xi) \qquad \text{for all } \eta_1, \eta_2, \xi \in V \text{ and } \lambda \in \mathbb{C}.$$

(1) Prove that every sesquilinear form satisfies the *polarization identity*

$$4(\eta|\xi) = \sum_{k=0}^{3} i^k(\xi + i^k\eta|\xi + i^k\eta). \tag{1.3.4}$$

(2) A sesquilinear form is called *positive* if $(\eta|\eta) \geq 0$ for all $\eta \in V$ and *self-adjoint* (or *Hermitian*) if $(\eta|\xi) = \overline{(\xi|\eta)}$ for all $\eta, \xi \in V$. Prove that positive implies self-adjoint.

Theorem 1.3.5 (Cauchy-Schwarz Inequality) — Suppose $\langle\cdot|\cdot\rangle$ is a positive sesquilinear form on a vector space H, and set $\|\xi\| = \sqrt{\langle\xi|\xi\rangle}$. We have

$$|\langle\eta|\xi\rangle| \leq \|\eta\| \cdot \|\xi\| \quad \text{for all } \eta, \xi \in H. \tag{C-S}$$

If moreover $\langle\cdot|\cdot\rangle$ is definite so that H is a Hilbert space, we have equality if and only if η, ξ are proportional.

Proof. We may assume $\langle\eta|\xi\rangle \in \mathbb{R}$ by multiplying η by a phase. Consider the real non-negative polynomial

$$p(t) := \langle\eta - t\xi|\eta - t\xi\rangle = \|\eta\|^2 - 2t\langle\eta|\xi\rangle + t^2\|\xi\|^2.$$

If $\|\xi\| = 0$, then the only way $p(t) \geq 0$ for all t is if $\langle\eta|\xi\rangle = 0$. If $\|\xi\| \neq 0$, then $p(t)$ achieves its minimum at $t_0 = \langle\eta|\xi\rangle/\|\xi\|^2$, at which

$$0 \leq p(t_0) = \|\eta\|^2 - \langle\eta|\xi\rangle^2/\|\xi\|^2 \qquad \Longleftrightarrow \qquad \langle\eta|\xi\rangle \leq \|\eta\| \cdot \|\xi\|.$$

Equality holds if and only if $p(t_0) = 0$. If $\langle\cdot|\cdot\rangle$ is definite, this is equivalent to $\eta = t_0\xi$. $\square$

Exercise 1.3.6. Deduce from the proof of the Cauchy-Schwarz Inequality (C-S) that if $\langle\cdot|\cdot\rangle$ is a positive sesquilinear form on a vector space H, then $\langle\xi|\xi\rangle = 0$ if and only if $\langle\eta|\xi\rangle = 0$ for all $\eta \in H$.

Exercise 1.3.7. Use the Cauchy-Schwarz Inequality (C-S) to show that $\|\cdot\|$ is a norm on H.

Exercise 1.3.8. Let $(V, \|\cdot\|)$ be a normed space. Show that the expression

$$\langle\eta|\xi\rangle := \frac{1}{4}\sum_{k=0}^{3} i^k\|\xi + i^k\eta\|^2$$

determines an inner product if and only if $\|\cdot\|$ satisfies the *parallelogram law*:

$$2\|\xi\|^2 + 2\|\eta\|^2 = \|\xi + \eta\|^2 + \|\xi - \eta\|^2.$$

Thus being a Hilbert space is a property of a normed space.

Definition 1.3.9 — An *orthonormal basis* (ONB) for H is a finite set $\{e_j\}_{j=1}^n \subset H$ such that

- (linearly independent) $\sum \lambda_j e_j = 0$ implies $\lambda_j = 0$ for all j,
- (spans) every $\eta \in H$ can be written as a linear combination $\eta = \sum \lambda_j e_j$ for some scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, and
- (orthonormal) $\langle e_i | e_j \rangle = \delta_{i=j} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else.} \end{cases}$

Each element e_i of an ONB is a *unit vector*, meaning it has unit length.

Example 1.3.10 — The *computational basis* for $\mathbb{C}^n$ is $\{|0\rangle, \ldots, |n-1\rangle\}$, where $|i\rangle$ is the vector which is one in the $(i+1)$-st coordinate and zero in every other coordinate.

Exercise 1.3.11. If $\{e_j\}_{j=1}^n \subset H$ is an ONB, then $\eta = \sum_{j=1}^n \langle e_j | \eta \rangle e_j$ for all $\eta \in H$.

Proposition 1.3.12 — An ONB exists for every Hilbert space.

Proof. We use the Gram-Schmidt algorithm. We assume we have a basis $\{v_1, \ldots, v_n\}$ of H. Set $e_1 := v_1 / \|v_1\|$, and then inductively set $w_k := v_k - \sum_{j=1}^{k-1} \langle e_j | v_k \rangle e_j$ and $e_k := w_k / \|w_k\|$. □

Definition 1.3.13 — A linear map between Hilbert spaces is called an *isomorphism* if it is invertible and preserves the inner product, i.e., $\langle x\eta | x\xi \rangle_K = \langle \eta | \xi \rangle_H$ for all $\eta, \xi \in H$. (In §1.4 below, we will call such maps *unitary*.)

Warning 1.3.14 — The notion of isomorphism of Hilbert spaces is finer than the notion of isomorphism of the underlying vector spaces.

Example 1.3.15 — Suppose x is an invertible operator on H. Define a second inner product on H by $\langle \eta | \xi \rangle_x := \langle x\eta | x\xi \rangle$, and let H_x denote H with this second inner product. The operator $H_x \to H$ given by $\eta \mapsto x\eta$ is unitary.

Corollary 1.3.16 — Each Hilbert space H is isomorphic as a Hilbert space to $\mathbb{C}^n$ where $n = \dim(H)$.

Proof. Let $\{e_1, \ldots, e_n\}$ be an ONB for H. The coordinate map $[\,\cdot\,] : \sum_{j=1}^n \lambda_j e_j \mapsto (\lambda_j)_{j=1}^n$ is the desired isomorphism. One checks that this map is invertible and satisfies $\langle \eta | \xi \rangle_H = \langle [\eta] | [\xi] \rangle_{\mathbb{C}^n}$ for all $\eta, \xi \in H$. □

Definition 1.3.17 — Given a Hilbert space H, the *conjugate space* $\overline{H}$ is the set of symbols $\{\overline{\eta} \,|\, \eta \in H\}$ with vector space structure given by

$$\overline{\eta} + \overline{\xi} := \overline{\eta + \xi} \qquad \text{and} \qquad \lambda \cdot \overline{\eta} := \overline{\overline{\lambda} \cdot \eta}$$

and inner product given by $\langle \overline{\eta} | \overline{\xi} \rangle := \langle \xi | \eta \rangle$. We identify $\overline{\overline{H}} = H$ itself.

The *dual space* $H^\vee$ is the space of linear functionals $H \to \mathbb{C}$. For example, every $\eta \in H$ gives a linear functional $\langle \eta | : H \to \mathbb{C}$ by $\langle \eta | \xi := \langle \eta | \xi \rangle$. We endow $H^\vee$ with the operator norm

$$\left\| \langle \eta | \right\| := \sup_{\|\xi\|=1} |\langle \eta | \xi \rangle|.$$

As with finite dimensional vector spaces, there is a canonical linear isomorphism $H \cong H^{\vee\vee}$, which is unitary by the following theorem.

Theorem 1.3.18 (Riesz Representation) — The dual space $H^\vee$ is canonically isomorphic as a Hilbert space to $\overline{H}$ via the map $\overline{\eta} \mapsto \langle \eta |$.

Proof. First, $\langle \lambda\eta + \xi | = \overline{\lambda} \langle \eta | + \langle \xi |$, so this map is linear and thus well-defined. If $\langle \eta | = \langle 0 |$, then $\langle \eta | \xi \rangle = 0$ for all $\xi \in H$, so $\eta = 0$. Finally, if $f : H \to \mathbb{C}$ is a non-zero linear functional, pick an ONB $\{e_j\}_{j=1}^n$ of H, and observe that f is completely determined by $f(e_1), \ldots, f(e_n)$. It is readily checked that $f = \sum_{j=1}^n f(e_j) \langle e_j |$.

We now check the map $\overline{\eta} \mapsto \langle \eta |$ is isometric. Observe that for all $\xi \in H$ with $\|\xi\| \leq 1$,

$$|\langle \eta | \xi \rangle| \underset{\text{(C-S)}}{\leq} \|\eta\| \cdot \|\xi\| \qquad \underset{\text{(Exer. 1.2.9)}}{\Longrightarrow} \qquad \left\| \langle \eta | \right\| \leq \|\eta\| = \|\overline{\eta}\|.$$

Conversely, $\langle \eta | \eta \rangle = \|\eta\|^2$ implies that $\|\overline{\eta}\| = \|\eta\| \leq \left\| \langle \eta | \right\|$. $\square$

Definition 1.3.19 — Given $x : H \to K$, we define the *conjugate* $\overline{x} : \overline{H} \to \overline{K}$ by $\overline{x}(\overline{\eta}) := \overline{x\eta}$. Observe that $\overline{\mathrm{id}_H} = \mathrm{id}_{\overline{H}}$, $\overline{\overline{x}} = x$ under the identification $\overline{\overline{H}} = H$, and $\overline{x \circ y} = \overline{x} \circ \overline{y}$ whenever x, y are composable.

Similarly, we define the *transpose* $x^\vee : K^\vee \to H^\vee$ by $\langle \xi | \mapsto \langle \xi | \circ x : H \to \mathbb{C}$. Observe that $(\mathrm{id}_H)^\vee = \mathrm{id}_{H^\vee}$, $x^{\vee\vee} = x$ under the identification $H^{\vee\vee} = H$, and $(x \circ y)^\vee = y^\vee \circ x^\vee$ whenever x, y are composable.

Exercise 1.3.20. Prove that the operations $(\cdot)^\dagger, \overline{(\cdot)}, (\cdot)^\vee$ all commute. Then prove that for $x : H \to K$, $\overline{x^\dagger} = x^\vee$.

Warning 1.3.21 — In order to identify the operation $x \mapsto x^\vee$ with the transpose, one must choose bases for H, K and use these bases to construct (highly non-canonical!) isomorphisms $H \cong H^\vee$ and $K \cong K^\vee$. Thus a more mathematically precise moniker for $x \mapsto x^\vee$ is the *Banach adjoint*. We obviously avoid this terminology to not confuse the reader with the Hilber space adjoint.

Exercise 1.3.22. Use the Riesz Representation Theorem 1.3.18 to show that the data of a Hilbert space is equivalent to a vector space V equipped with an isomorphism $\varphi\colon \overline{V} \cong V^\vee := \mathrm{Hom}(V \to \mathbb{C})$ such that:

- (Hermitian) For every $\xi, \eta \in V$, $\varphi(\overline{\xi})(\eta) = \overline{\varphi(\overline{\eta})(\xi)}$, and
- (Positive) For every $\xi \in V$, $\varphi(\overline{\xi})(\xi) \geq 0$ in $\mathbb{C}$.

1.4 Operators

Linear maps between Hilbert spaces are also commonly called *(linear) operators*. Given Hilbert spaces H, K, we denote the space of linear operators $H \to K$ by $B(H \to K)$ or $\mathrm{Hom}(H \to K)$, and we write $B(H) = B(H \to H)$ and $\mathrm{End}(H) = \mathrm{Hom}(H \to H)$.

Definition 1.4.1 — Given a linear map $x : H \to K$, observe that the map

$$H \ni \eta \mapsto \langle \xi | x\eta \rangle \quad \text{for } \xi \in K$$

is a linear functional in $H^\vee$. Hence there is some vector $x^\dagger \xi \in H$ such that the above map is equal to $\langle x^\dagger \xi|$. It can be verified that the *adjoint map*

$$\begin{aligned} \dagger\colon \overline{B(H \to K)} &\to B(K \to H) \\ x &\mapsto x^\dagger \end{aligned}$$

is an isomorphism such that:

- When x, y are composable operators, $(xy)^\dagger = y^\dagger x^\dagger$, and
- $x^{\dagger\dagger} = x$ for all operators x.

We call $x^\dagger$ the *adjoint* of x.

Just as we defined coordinates for linear maps between finite dimensional vector spaces, we have coordinates for operators between finite dimensional Hilbert spaces.

Definition 1.4.2 (Orthonormal coordinates) — Given ONBs $\{e_j\}_{j=1}^n$ for H and $\{f_k\}_{k=1}^m$

for K, the coordinate isomorphism $B(H \to K) \cong M_{m\times n}(\mathbb{C})$ is given by

$$x \longmapsto [x] \coloneqq (\langle f_i | x e_j \rangle)_{i,j}. \tag{1.4.3}$$

Here, the columns correspond to the ONB of the source, and the rows correspond to the ONB of the target. In addition to composition of linear operators corresponding to matrix multiplication, since we used ONBs for our coordinates, we also have that the adjoint corresponds to the conjugate transpose, also denoted by $\dagger$. That is, the following diagram commutes.

$$\begin{tikzcd}
(K, \{f_k\}_{k=1}^m) \arrow[r, "x^\dagger"] \arrow[d, "{[\,\cdot\,]}"'] & (H, \{e_j\}_{j=1}^n) \arrow[d, "{[\,\cdot\,]}"] \\
\mathbb{C}^m \arrow[r, "{[x]^\dagger}"'] & \mathbb{C}^n
\end{tikzcd}$$

Thus studying operators between Hilbert spaces and operations between them reduces to studying matrices and their operations. Below, we use the abstract language of Hilbert spaces, but the reader may safely replace $B(H)$ with $M_n(\mathbb{C})$ if they choose.

Exercise 1.4.4. Suppose $B \subset H$ is a basis which is not an ONB. Find an operator $x \in B(H)$ such that $[x^\dagger] \neq [x]^\dagger$.

Exercise 1.4.5 (Bra-ket notation)**.** Every vector η in a Hilbert space H gives rise to its associated *bra* operator $\langle \eta | \colon H \to \mathbb{C}$ given by $\xi \mapsto \langle \eta | \xi \rangle$. Similarly, there is an associated *ket* operator $|\eta\rangle \colon \mathbb{C} \to H$ given by $\lambda \mapsto \lambda\eta$.

(1) Prove that the map $\eta \mapsto |\eta\rangle$ is an isomorphism of vector spaces $H \cong \operatorname{Hom}(\mathbb{C} \to H)$.

(2) Show $\langle \eta |^\dagger = |\eta\rangle$ for $\eta \in H$.

(3) Interpret the unitary $\overline{H} \to H^\vee$ from the Riesz Representation Theorem 1.3.18 as the adjoint $\dagger : \overline{B(\mathbb{C} \to H)} \to B(H \to \mathbb{C})$.

(4) For $\eta, \xi \in H$, verify $\langle \eta | \xi \rangle = \langle \eta | \circ | \xi \rangle \in B(\mathbb{C}) \cong \mathbb{C}$.[1]

(5) For an ONB $\{e_j\}_{j=1}^n$ of H, prove $1_H = \sum |e_j\rangle\langle e_j|$ and deduce $f = \sum f(e_j)\langle e_j|$ for $f \in H^\vee$.

Exercise 1.4.6. Show the *polarization identity* (1.3.4) for operators:

$$4x^\dagger y = \sum_{k=0}^{3} i^k (x + i^k y)^\dagger (x + i^k y).$$

[1]That is, a "bra" and a "ket" form a "braket", motivating this terminology coming from physics.

Definition 1.4.7 — Given $\eta \in H$ and $\xi \in K$, the *rank one* operator $|\xi\rangle\langle\eta| : H \to K$ is given by $\zeta \mapsto \langle\eta|\zeta\rangle\xi$.

Example 1.4.8 — If $x \in M_n(\mathbb{C})$ commutes with all $y \in M_n(\mathbb{C})$, then $x = \lambda 1$ for some $\lambda \in \mathbb{C}$. To see this, observe that since $x|e_i\rangle\langle e_j| = |e_i\rangle\langle e_j|x$ for all i, j, it follows that x is diagonal, and all diagonal entries are equal.

Definition 1.4.9 — An operator $x \in B(H)$ is called:

- *normal* if $xx^\dagger = x^\dagger x$
- *self-adjoint* (or *Hermitian*) if $x^\dagger = x$
- *positive* (or *non-negative*) if $\langle\xi|x\xi\rangle \geq 0$ for all $\xi \in H$ (denoted $x \geq 0$)
- an (orthogonal) *projection* if $x^2 = x = x^\dagger$

An operator $u \in B(H \to K)$ is called:

- a *partial isometry* if $u^\dagger u$ is a projection.
- an *isometry* if $u^\dagger u = 1$.
- a *unitary* if u is invertible with $u^{-1} = u^\dagger$.

Remark 1.4.10. Here we have the following implications:

$$\begin{array}{ccccccc} & & & & \text{unitary} & \Rightarrow \text{isometry} \Rightarrow & \text{partial isometry} \\ & & & & \Downarrow & & \\ \text{projection} & \Rightarrow & \text{positive} & \Rightarrow \text{self-adjoint} \Rightarrow & \text{normal} & & \end{array}$$

We will show that positive implies self-adjoint in Corollary 1.4.19.

Example 1.4.11 — Given an orthonormal set $S = \{e_1, \ldots, e_k\} \subset H$, we get an orthogonal projection onto $\operatorname{span}(S)$ by $\sum_{j=1}^k |e_j\rangle\langle e_j|$.

Exercise 1.4.12. Given an orthonormal set $\{e_1, \ldots, e_k\} \subset H$ and another orthonormal set $\{f_1, \ldots, f_k\} \subset K$ of the same size, prove that $\sum_{j=1}^k |e_j\rangle\langle f_j|$ is a partial isometry.

Exercise 1.4.13. Prove that the projection $\sum_{j=1}^k |e_j\rangle\langle e_j|$ is independent of the choice of ONB of $\operatorname{span}(S)$.

Definition 1.4.14 — A *system of matrix units* in $B(H)$ is a collection of operators $\{e_{ij}\}$ satisfying

(SMU1) $e_{ij}e_{k\ell} = \delta_{j=k}e_{i\ell}$,

(SMU2) $\sum_j e_{jj} = 1$, and

(SMU3) $e_{ij}^\dagger = e_{ji}$

for all i, j, k, ℓ. Observe that each e_{ij} is a partial isometry and each e_{ii} is an orthogonal projection.

Exercise 1.4.15. Let $\{e_i\}$ be an ONB for H. Show that $e_{ij} := |e_i\rangle\langle e_j|$ is a SMU in $B(H)$.

Exercise 1.4.16. What sizes of systems of matrix units can occur in $M_n(\mathbb{C})$?

Lemma 1.4.17 — For all $x \in M_{m\times n}(\mathbb{C})$, $\ker(x) = \ker(x^\dagger x)$. In particular, $x^\dagger x = 0$ implies $x = 0$.

Proof. Clearly $x\eta = 0$ implies $x^\dagger x\eta = 0$. Conversely, if $x^\dagger x\eta = 0$, then $\|x\eta\|^2 = \langle \eta | x^\dagger x \eta\rangle = 0$, so $x\eta = 0$. For the final statement, observe $x^\dagger x = 0$ if and only if $\ker(x) = \ker(x^\dagger x) = \mathbb{C}^n$. □

Lemma 1.4.18 (Vector states separate points) — The following are equivalent for $x \in B(H)$.

(1) $x = 0$,

(2) $\langle \eta | x\eta\rangle = 0$ for all $\eta \in H$, and

(3) $\langle \eta | x\xi\rangle = 0$ for all $\eta, \xi \in H$.

Proof. Clearly (1) ⇒ (2).

(2)⇒(3): Suppose $\langle \eta | x\eta\rangle = 0$ for all $\eta \in H$. Consider the sesquilinear form $(\eta|\xi) := \langle \eta | x\xi\rangle$. By (1.3.4),

$$4\langle \eta | x\xi\rangle = 4(\eta|\xi) = \sum_{k=0}^{3} i^k(\xi + i^k\eta | \xi + i^k\eta) = \sum_{k=0}^{3} i^k\langle \xi + i^k\eta | x(\xi + i^k\eta)\rangle = 0 \qquad \forall\, \eta, \xi \in H.$$

Thus $\langle \eta | x\xi\rangle = 0$ for all $\eta, \xi \in H$.

(3)⇒(1): If $\langle \eta | x\xi\rangle = 0$ for all $\eta, \xi \in H$, then $x\xi = 0$ for all $\xi \in H$, so $x = 0$. □

Corollary 1.4.19 — Positive operators are self-adjoint.

Proof. When $x \geq 0$,

$$\langle \xi | x^\dagger \xi \rangle = \overline{\langle x^\dagger \xi | \xi \rangle} = \overline{\langle \xi | x \xi \rangle} = \langle \xi | x \xi \rangle \geq 0 \qquad \text{for all } \xi \in H.$$

Hence $\langle \xi | (x - x^\dagger)\xi \rangle = 0$ for all $\xi \in H$, and thus $x = x^\dagger$ by Lemma 1.4.18. □

Exercise 1.4.20. Prove that $x \in B(H)$ is normal if and only if $\|x\eta\| = \|x^\dagger \eta\|$ for all $\eta \in H$.

Exercise 1.4.21. Show that the map $x \mapsto \langle \cdot | x \cdot \rangle$ is a bijective correspondence between $B(H)$ and sesquilinear forms on H. Then prove that $\langle \cdot | x \cdot \rangle$ is self-adjoint/positive if and only if x is self-adjoint/positive respectively. Finally, show $\langle \cdot | x \cdot \rangle$ is *non-degenerate* (for every $\eta \in H$, there is a $\xi \in H$ with $\langle \eta | x \xi \rangle \neq 0$) if and only if x is invertible.
Hint: For a graphical calculus proof, see Corollary 1.6.21.

Exercise 1.4.22. We say two projections $p, q \in B(H)$ are *(Murray-von Neumann) equivalent*, denoted $p \approx q$, if there is a partial isometry $u \in B(H)$ such that $uu^* = p$ and $u^*u = q$. Prove that $\approx$ is an equivalence relation on $P(B(H))$, the set of projections of $B(H)$. Then describe the set of equivalence classes $P(M_n(\mathbb{C}))/\approx$.

Proposition 1.4.23 (Subspace-projection correspondence) — There is a bijective correspondence between orthogonal projections in $B(H)$ and subspaces of H given by $p \mapsto pH$ and $H \supset K \mapsto \sum_{j=1}^{k} |e_j\rangle\langle e_j|$ where $\{e_j\}_{j=1}^{k}$ is an ONB of K.

Proof. Note that the second map is well-defined by Exercise 1.4.13. Given an ONB $\{e_j\}_{j=1}^{k}$ of K, $\operatorname{im}\left(\sum_{j=1}^{k} |e_j\rangle\langle e_j|\right) = K$. Given a projection p and an ONB $\{e_j\}_{j=1}^{k}$ of pH, notice Exercise 1.3.11 yields

$$p\eta = \sum_{j=1}^{k} \langle e_j | p\eta \rangle e_j = \sum_{j=1}^{k} \langle pe_j | \eta \rangle e_j = \sum_{j=1}^{k} \langle e_j | \eta \rangle e_j = \sum_{j=1}^{k} |e_j\rangle\langle e_j | \eta.$$ □

Exercise 1.4.24. For a subspace $K \subset H$, its *orthogonal complement* is

$$K^\perp := \{\eta \in H \mid \langle \xi | \eta \rangle = 0 \text{ for all } \xi \in K\}.$$

(1) Show that if p is the orthogonal projection onto K, then $1 - p$ is the orthogonal projection onto $K^\perp$.

(2) Show that $H = K \oplus K^\perp$ is the *internal orthogonal direct sum*. That is, $K \cap K^\perp = 0$, vectors in K and $K^\perp$ are orthogonal, and $H = K + K^\perp$.

Definition 1.4.25 — Let $K \subseteq H$ be a subspace of H and $x \in B(H)$. We say K is *invariant* for x when $xK \subseteq K$.

Corollary 1.4.26 — Suppose $p \in B(H)$ is a projection and $x \in B(H)$ is an operator.

(1) pH is invariant for x if and only if $xp = pxp$.

(2) pH is invariant for x and $x^\dagger$ if and only if $xp = px$.

Proof. To prove (1), observe that pH invariant for x means that $pxp\eta = xp\eta$ for all $\eta \in H$, and thus $pxp = xp$. Conversely, if $pxp = xp$, then $xpH \subseteq pH$.

To prove (2), observe that if pH invariant for x and $x^\dagger$ is equivalent to $xp = pxp$ and $x^\dagger p = px^\dagger p$. Hence $px = (x^\dagger p)^\dagger = (px^\dagger p)^\dagger = pxp = xp$. Conversely, $xp = px$ implies both $xp = pxp$ and $x^\dagger p = px^\dagger p$. ☐

Remark 1.4.27. Given an operator $x \in B(H)$ and a projection $p \in B(H)$, we can view x as matrix with operator entries

$$x = \begin{bmatrix} pxp & px(1-p) \\ (1-p)xp & (1-p)x(1-p) \end{bmatrix}$$

acting on $pH \oplus (1-p)H = H$. Under this identification,

$$p = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{and} \qquad 1 - p = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus (1) of Corollary 1.4.26 above is equivalent to x being upper triangular, and (2) of Corollary 1.4.26 above is equivalent to x being diagonal.

Proposition 1.4.28 — The following are equivalent for $u \in B(H)$:

(1) u is a partial isometry.

(2) $u = uu^\dagger u$.

(3) $u^\dagger = u^\dagger u u^\dagger$.

(4) $u^\dagger$ is a partial isometry.

Proof.

(1) ⇒ (2): Observe that

$$(u - uu^\dagger u)^\dagger (u - uu^\dagger u) = (u^\dagger - u^\dagger uu^\dagger)(u - uu^\dagger u) = u^\dagger u - 2u^\dagger uu^\dagger u + u^\dagger uu^\dagger uu^\dagger u = 0,$$

so $u - uu^\dagger u = 0$ by Lemma 1.4.17.

(2) $\Rightarrow$ (1): Multiply both sides on the left by $u^\dagger$ to see $u^\dagger u = u^\dagger u u^\dagger u$, which is self-adjoint.

(2) $\Leftrightarrow$ (3): Just take adjoints.

(3) $\Leftrightarrow$ (4): Apply (1) $\Leftrightarrow$ (2) to $u^\dagger$. □

Remark 1.4.29. Proposition 1.4.28 above gives a geometric interpretation about what it means to be a partial isometry. Indeed, $u : H \to K$ restricts to a unitary isomorphism from $u^\dagger uH$ onto $uu^\dagger K$ with inverse $u^\dagger$. Every partial isometry is of this form; that is, given projections $p \in B(H)$ and $q \in B(K)$ and a unitary isomorphism $u : pH \to qK$, we can extend u to an operator in $B(H \to K)$ satisfying $u^\dagger u = p$ and $uu^\dagger = q$ by defining u on $(1-p)H$ to be zero. Viewing $u\colon pH \oplus (1-p)H \to qH \oplus (1-q)H$ as in Remark 1.4.27,

$$u = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Definition 1.4.30 — Define a partial order on $B(H)$ by $x \leq y$ if $y - x \geq 0$.

Exercise 1.4.31. Show that if $x \leq y$ and $z \in B(H)$, then $z^\dagger x z \leq z^\dagger y z$.

Exercise 1.4.32. Prove that for projections, $p \leq q$ if and only if $pH \subseteq qH$.

Exercise 1.4.33. Show that if $p_1, \ldots, p_n$ are projections such that $\sum p_j = 1$, then $p_i p_j = 0$ when $i \neq j$. Deduce that (SMU1) can be replaced with $e_{ij}e_{jk} = e_{ik}$ for all i, j, k.

Definition 1.4.34 — We call a non-zero projection $p \in B(H)$ *minimal* if $pB(H)p = \mathbb{C}p$.

Exercise 1.4.35. Show that the following are equivalent for a non-zero projection p.

(1) p is minimal.

(2) $0 \leq q \leq p$ implies $q = 0$ or $q = p$.

(3) pH is 1-dimensional ($\operatorname{rank}(p) = 1$).

(4) $p = |\xi\rangle\langle\xi|$ for some unit vector $\xi \in H$.

1.5 Direct sum and tensor product

Definition 1.5.1 — Given two Hilbert spaces H, K, their *direct sum* is defined as a Hilbert space $H \oplus K$ together with isometries $i_H : H \to H \oplus K$ and $i_K : K \to H \oplus K$ which satisfy $i_H i_H^\dagger + i_K i_K^\dagger = 1$. By Proposition 1.4.28 and Exercise 1.4.33, it follows that $i_H^\dagger i_K = 0$ and $i_K^\dagger i_H = 0$. By Remark 1.4.27, operators on $H \oplus K$ can be viewed as matrices of operators:

$$x = \begin{bmatrix} i_H^\dagger x i_H & i_H^\dagger x i_K \\ i_K^\dagger x i_H & i_K^\dagger x i_H \end{bmatrix} \in \begin{bmatrix} B(H \to H) & B(K \to H) \\ B(H \to K) & B(K \to K) \end{bmatrix}.$$

Given a second direct sum $H \oplus' K$ with isometries $j_H : H \to H \oplus' K$ and $j_K : K \to H \oplus' K$ satisfying $j_H j_H^\dagger + j_K j_K^\dagger = 1$, there is a unique unitary isomorphism $u := j_H i_H^\dagger + j_K i_K^\dagger : H \oplus K \to H \oplus' K$ which is compatible with i_H, i_K and j_H, j_K in the sense that the following diagram commutes.

$$\begin{tikzcd} H \arrow[r, "i_H"] \arrow[dr, "j_H"'] & H \oplus K \arrow[d, "u"] & K \arrow[l, "i_K"'] \arrow[dl, "j_K"] \\ & H \oplus' K & \end{tikzcd}$$

The uniqueness of the above unitary implies that given three choices for the direct sum $H \oplus K, H \oplus' K, H \oplus'' K$, the unique unitaries fit into the following commutative diagram.

$$\begin{tikzcd} H \oplus K \arrow[rr, "\cong"] \arrow[dr, "\cong"'] & & H \oplus'' K \\ & H \oplus' K \arrow[ur, "\cong"'] & \end{tikzcd}$$

Thus models for $H \oplus K$ form a *contractible space*.

Ethos 1.5.2 — There is an important distinction between a *property* and a *structure* for mathematical objects. To highlight this difference we consider the following two mathematical objects:

- the direct sum $H \oplus K$ of two Hilbert spaces, and
- an ONB for a Hilbert space H.

Here, we write *the* direct sum, as the space of direct sums is contractible. Even though a choice of direct sum comes equipped with choices of isometries, these maps are completely determined by a universal property, and there is a unique isomorphism between any two choices. Thus choosing a model for the direct sum is not a substantive choice in any way.

On the other hand, we write *an* ONB, indicating this choice is much more significant. Indeed, there is a set of ONBs for any Hilbert space; for $\mathbb{C}^n$, this set can be identified with the *unitary group*

$$U(n) = \left\{ u \in M_n(\mathbb{C}) \,\middle|\, uu^\dagger = u^\dagger u = 1 \right\}.$$

There are no 'maps' between ONBs.

On a similar note, a vector space is not a Hilbert space, as there is a set of possible inner products. For the vector space $\mathbb{C}$, this set can be identified with $\mathbb{R}_{>0} = [0, \infty)$. We warn the reader that even though $\mathbb{R}_{>0}$ is contractible as a topological space with the *norm topology*, the space of inner products is really $\mathbb{R}_{>0}$ with the *discrete topology*.

Exercise 1.5.3. Prove that an internal orthogonal direct sum decomposition $H = K \oplus K^\perp$ as in Exercise 1.4.24 is a direct sum of Hilbert spaces.

Exercise 1.5.4. Show that Definition 1.5.1 agrees with the usual definition, i.e.,

$$H \oplus K := \{(\eta, \xi) \mid \eta \in H, \xi \in K\} \qquad \langle(\eta_1, \xi_1)|(\eta_2, \xi_2)\rangle := \langle\eta_1|\eta_2\rangle_H + \langle\xi_1|\xi_2\rangle_K$$

with the obvious inclusions satisfies the universal property.

Exercise 1.5.5 (Quotient spaces)**.** Suppose V is a vector space and $W \subset V$ is a subspace. Define V/W as the set whose elements are symbols of the form $v + W$ for $v \in V$ up to the equivalence relation $u + W = v + W$ iff $u - v \in W$.

(1) Show that V/W is a vector space with operations

$$\begin{aligned} (u + W) + (v + W) &:= (u + v) + W & \qquad u, v &\in V \\ \lambda \cdot (u + W) &:= (\lambda u) + W & \qquad \lambda \in \mathbb{C},\ u &\in V. \end{aligned}$$

(2) Observe that by definition, the *canonical surjection* $V \to V/W$ by $v \mapsto v + W$ is linear with kernel W. Show V/W satisfies the following universal property:

- For every map $f : V \to U$ into a vector space U with $W \subseteq \ker f$, there exists a unique map $\widetilde{f} : V/W \to U$ such that the following diagram commutes:

$$\begin{tikzcd} 0 \arrow[r] & W \arrow[r] \arrow[dr, "0"'] & V \arrow[r] \arrow[d, "f"] & V/W \arrow[r] \arrow[dl, dashed, "\exists! \widetilde{f}"] & 0 \\ & & U & & \end{tikzcd}$$

Now consider when $K \subset H$ are Hilbert spaces.

(3) Show that the map $H/K \to [0, \infty)$ given by $\|\eta + K\| := \inf_{\xi \in K} \|\eta + \xi\|$ is a norm.

(4) Prove that the map $K^\perp \to H/K$ given by $\eta \mapsto \eta + K$ is isometric and invertible. Deduce that the infemum norm above is a Hilbert space norm on H/K.

(5) Deduce the analogous universal property for Hilbert space quotients:

- For every operator $x\colon H \to L$ into a Hilbert space L with $K \subseteq \ker x$, there exists a unique operator $\widetilde{x}\colon H/K \to L$ such that the following diagram commutes:

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r] \arrow[dr, "0"'] & H \arrow[r] \arrow[d, "x"] & H/K \arrow[r] \arrow[dl, dashed, "\exists!\widetilde{x}"] & 0 \\
& & L & &
\end{tikzcd}$$

Definition 1.5.6 — Given two Hilbert spaces H, K, their *tensor product* is the Hilbert space $H \otimes K$, which can be defined in a number of ways. The easiest is in terms of choosing ONBs $\{e_j\}$ of H and $\{f_k\}$ of K. The tensor product $H \otimes K$ then has ONB the formal symbols $\{e_j \otimes f_k\}$. Thus $\dim(H \otimes K) = \dim(H) \otimes \dim(K)$.

It can be readily checked that if we chose different ONBs $\{e'_j\}$ of H and $\{f'_k\}$ of K, there is a canonical isomorphism of Hilbert spaces from the Hilbert space $H \otimes K$ with ONB $\{e_j \otimes f_k\}$ to the Hilbert space $H \otimes' K$with ONB $\{e'_j \otimes f'_k\}$. This map is canonical in the sense that given a third choice $\{e''_j\}$ of H and $\{f''_k\}$ of K, the canonical isomorphisms between $H \otimes K, H \otimes' K, H \otimes'' K$ fit in the following commutative diagram.

$$\begin{tikzcd}
H \otimes K \arrow[rr, "\cong"] \arrow[dr, "\cong"'] & & H \otimes'' K \\
& H \otimes' K \arrow[ur, "\cong"'] &
\end{tikzcd}$$

Thus models for $H \otimes K$ form a *contractible space*.

Example 1.5.7 — The computational basis for $\mathbb{C}^m \otimes \mathbb{C}^n$ is usually denoted by

$$\{|ij\rangle \mid i = 0, \ldots, m-1 \text{ and } j = 0, \ldots, n-1\}.$$

Exercise 1.5.8. Show that $H \otimes K \cong \bigoplus_{i=1}^{\dim(H)} K$ and $H \otimes K \cong \bigoplus_{j=1}^{\dim(K)} H$.

Definition 1.5.9 — The *Kronecker product* of $x \in M_{k\times\ell}(\mathbb{C})$ and $y \in M_{m\times n}(\mathbb{C})$ is the matrix

$$x \otimes y := \begin{pmatrix} x_{11}y & \cdots & x_{1\ell}y \\ \vdots & & \vdots \\ x_{k1}y & \cdots & x_{k\ell}y \end{pmatrix} \in M_{km\times\ell n}(\mathbb{C}).$$

Exercise 1.5.10. Show that the choices of ONBs for H, K give a canonical isomorphism $B(H \otimes K) \cong B(H) \otimes B(K)$.

One can also define the tensor product via *universal property*, representing *bilinear* maps (which are linear in both entries). Any object which satisfies the universal property is unique up to unique isomorphism, i.e. the space of such objects is contractible.

Definition 1.5.11 — The tensor product Hilbert space of H, K is a Hilbert space $H \otimes K$ together with a bilinear map $\otimes : H \times K \to H \otimes K$ which satisfy the universal property that for every Hilbert space L and every *bilinear* map $T : H \times K \to L$, there is a unique *linear* map $\widetilde{T} : H \otimes K \to L$ such that the following diagram commutes.

$$\begin{tikzcd}
H \otimes K \arrow[rr, "T"] \arrow[dr, "\otimes"'] & & L \\
& H \otimes K \arrow[ur, "\widetilde{T}"'] &
\end{tikzcd}$$

Exercise 1.5.12. Use the universal property above to prove that the tensor product Hilbert space $(H \otimes K, \otimes : H \times K \to H \otimes K)$ is unique up to unique isomorphism.

Exercise 1.5.13. Show that Definition 1.5.6 and Definition 1.5.11 agree, i.e., the construction in Definition 1.5.6 satisfies the universal property in Definition 1.5.11.

Exercise 1.5.14. Construct a linear isomorphism $H \otimes \overline{K} \cong B(K \to H)$.

Exercise 1.5.15. Suppose H is a Hilbert space and $\{e_i\}$ is an ONB. Show that the element $\sum e_i \otimes \overline{e_i} \in H \otimes \overline{H}$ is independent of the choice of ONB.

Exercise 1.5.16 (Associator)**.** Suppose H_1, H_2, H_3 are Hilbert spaces. Construct *associator* unitary isomorphisms

$$\alpha_{H_1 H_2 H_3} : H_1 \otimes (H_2 \otimes H_3) \xrightarrow{\sim} (H_1 \otimes H_2) \otimes H_3.$$

Do this in two ways: by hand and using Exercise 1.5.12. Then prove that the following diagram commutes:

$$\begin{tikzcd}
& (H_1 \otimes H_2) \otimes (H_3 \otimes H_4) \arrow[dr, "\alpha"] & \\
H_1 \otimes (H_2 \otimes (H_3 \otimes H_4)) \arrow[ur, "\alpha"] \arrow[d, "1 \otimes \alpha"] & & ((H_1 \otimes H_2) \otimes H_3) \otimes H_4 \\
H_1 \otimes ((H_2 \otimes H_3) \otimes H_4) \arrow[rr, "\alpha"'] & & (H_1 \otimes (H_2 \otimes H_3)) \otimes H_4 \arrow[u, "\alpha \otimes 1"']
\end{tikzcd}$$

Exercise 1.5.17 (Parenthetizations)**.** Use induction to show that the number of ways of parenthesizing a word of length n is given by the n-th Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

For example, there are 5 ways to fully parenthesize a word of length 4, corresponding to the vertices of the pentagon in Exercise 1.5.16 above. Then show that C_n satisfies the following recurrence relation:

$$C_n = \sum_{k=0}^{n-1} C_k C_{n-k} \qquad \text{with } C_0 = 1.$$

Exercise 1.5.18 (Unitors)**.** Construct left and right *unitor* unitary isomorphisms

$$\lambda_H : \mathbb{C} \otimes H \xrightarrow{\sim} H \xleftarrow{\sim} H \otimes \mathbb{C} : \rho_H.$$

Do this in two ways: by hand and using Exercise 1.5.12. Then prove that the following diagram commutes:

$$\begin{tikzcd}
H \otimes (\mathbb{C} \otimes K) \arrow[rr, "\alpha"] \arrow[dr, "1\otimes\rho"'] & & (H \otimes \mathbb{C}) \otimes K \arrow[dl, "\lambda\otimes 1"] \\
& H \otimes K &
\end{tikzcd}$$

Exercise 1.5.19. Construct *braiding* unitary isomorphisms

$$\beta_{HK} \colon H \otimes K \xrightarrow{\sim} K \otimes H.$$

Do this in two ways: by hand and using Exercise 1.5.12. Then prove that the following diagram commutes:

$$\begin{tikzcd}
& H_3 \otimes (H_1 \otimes H_2) \arrow[r, "\alpha"] & (H_3 \otimes H_1) \otimes H_2 & \\
(H_1 \otimes H_2) \otimes H_3 \arrow[ur, "\beta"] & & & (H_1 \otimes H_3) \otimes H_2 \arrow[ul, "\beta\otimes 1"'] \\
& H_1 \otimes (H_2 \otimes H_3) \arrow[ul, "\alpha"] \arrow[r, "1\otimes\beta"'] & H_1 \otimes (H_3 \otimes H_2) \arrow[ur, "\alpha"'] &
\end{tikzcd}$$

Note: This is known as the Yang-Baxter equation.

Furthermore, show that the following symmetry condition holds:

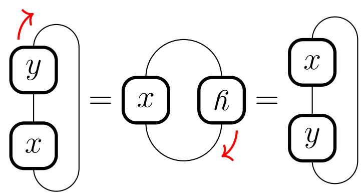
$$\begin{tikzcd}
H \otimes K \arrow[rr, "1"] \arrow[dr, "\beta"'] & & H \otimes K \\
& K \otimes H \arrow[ur, "\beta"'] &
\end{tikzcd}$$

1.6 Graphical calculus

In this section, we introduce a powerful and elegant *graphical calculus* for working with operators between Hilbert spaces. When we first learn algebra, we write equations linearly, e.g.,

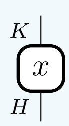
$$ax^2 + bx + c = 0.$$

Ironically, *linear* algebra is best described using equations drawn in a 2-dimensional *plane*, as we have two composition operations: composition of linear operators and tensor product. These two operations satisfy a certain commutation relation, allowing for this 2-dimensional graphical calculus. For example, the *trace* of a linear operator $a \in M_n(\mathbb{C})$ is given by the linear expression $\operatorname{Tr}(a) := \sum_{i=1}^{n} a_{ii}$. The fact that $\operatorname{Tr}(xy) = \operatorname{Tr}(yx)$ for $x \in M_{m\times n}(\mathbb{C})$ and $y \in M_{n\times m}(\mathbb{C})$, while following from a simple algebraic computation, is best proven via this 2-dimensional graphical calculus.

$\operatorname{Tr}(yx) =$  $= \operatorname{Tr}(xy)$ (1.6.1)

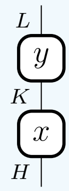
In this section, we build the necessary tools to understand how the above expression constitutes a proof.

Definition 1.6.2 — We represent the Hilbert space H by drawing a *line/string/strand* labeled by the space, and we represent an operator $x : H \to K$ by a coupon labeled by x between the two labeled strands. (We will sometimes omit labels if they can be inferred from the rest of the diagram.) By convention, we represent $\mathbb{C}$ by drawing the *empty strand*; see Warning 1.6.5 below. (We sometimes denote $\mathbb{C}$ with a dotted line to avoid confusion or to make a pedagogical point, but we try to represent it by empty space as much as possible.) We read our diagrams from *bottom-to-top*. The authors endearingly refer to this as the *optimistic convention* (always look up!).

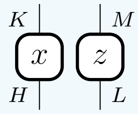


$\longleftrightarrow \qquad x : H \to K.$

The composite of $x : H \to K$ with $y : K \to L$ is denoted by the vertical stacking of diagrams. The tensor product of x with $z : L \to M$ is given by the horizontal juxtiposition of diagrams.



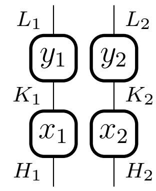
$\longleftrightarrow \qquad y \circ x \colon H \to L$



$\longleftrightarrow \qquad x \otimes z \colon H \otimes L \to K \otimes M$

There is an immediate ambiguity in interpreting the following diagram:

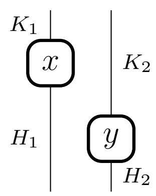
$$(y_1 \circ x_1) \otimes (y_2 \circ x_2) \stackrel{?}{=}$$



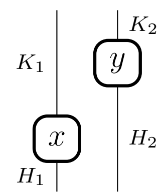
$$\stackrel{?}{=} (y_1 \otimes y_2) \circ (x_1 \otimes x_2). \tag{1.6.3}$$

Fortunately, these two expressions agree *on the nose*, so there is actually no such ambiguity.

Exercise 1.6.4. Prove the *interchange law*: for all $x : H_1 \to K_1$ and $y : H_2 \to K_2$,



$$= (x \otimes \mathrm{id}_{K_2}) \circ (\mathrm{id}_{H_1} \otimes y) = (\mathrm{id}_{K_1} \otimes y) \circ (x \otimes \mathrm{id}_{H_2}) =$$

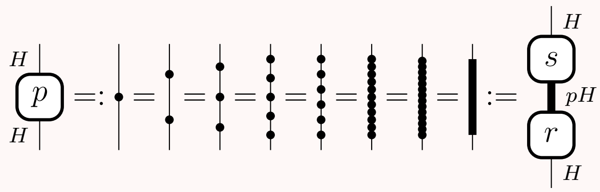


.

Warning 1.6.5 — We will rely on Exercise 1.5.16 to *completely ignore* the difference between $H \otimes (K \otimes L)$ and $(H \otimes K) \otimes L$. Similarly, we rely on Exercise 1.5.18 to *completely ignore* copies of $\mathbb{C}$ under tensor product and represent them as the empty line. We will discuss why this is OK later in Part [[II]] §[[?]].

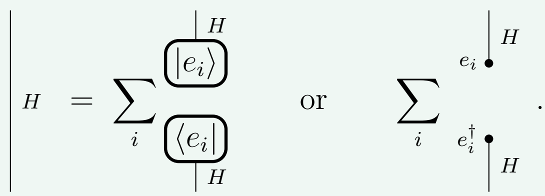
Ethos 1.6.6 — We now recontextualize the correspondence between projections and subspaces from Proposition 1.4.23 in terms of this diagrammatic calculus. Given a Hilbert space H, a projection $p \in B(H)$ satisfies $p = p^n$ for all n. So in the presence of a p on a H-string, we can proliferate copies of p until we think of them as densley populating a section of the H-strand. Taking some kind of *continuum/thermodynamic limit*, we can think of p as *splitting* as a *retract* $r : H \to pH$ and a *section* $s : pH \hookrightarrow H$

such that $r \circ s = \mathrm{id}_{pH}$ and $s \circ r = p$.



We will explore this notion in detail in Part [[II]] §[[?]].

Example 1.6.7 — We have the following important *skein relation* which may be applied *locally* in any string diagram. Given an ONB $\{e_i\}$ for H, the relation $\mathrm{id}_H = \sum_i |e_i\rangle\langle e_i|$ is represented graphically by



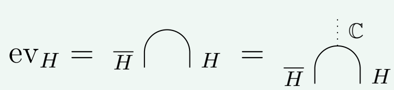
In this second graphical representation, we identify $e_i = |e_i\rangle : \mathbb{C} \to H$ and simply write $e_i^\dagger$ for $\langle e_i|$.

Warning 1.6.8 — In the graphical calculus for Hilbert spaces, we will elide the difference between $\overline{H}$ and $H^\vee$ as they are canonically unitarily isomorphic via the Riesz Representation Theorem 1.3.18. Often, we will choose the more compact notation $\overline{H}$ over the heavy notation $H^\vee$. However, it is important to note that the linear transpose operation $x \mapsto x^\vee$ and the anti-linear conjugation operation $x \mapsto \overline{x}$ are not the same on operators. (See Definition 1.3.19.) We will be carefuly to use the notation $H^\vee$ when we are discussing the transpose operation to avoid confusion.

Example 1.6.9 — There is a distinguished map $\mathrm{ev}_H : \overline{H} \otimes H \to \mathbb{C}$ given by *evaluation* of the covector $\langle \eta|$ at the vector $|\xi\rangle$:

$$\mathrm{ev}_H(\langle \eta| \otimes |\xi\rangle) := \langle \eta|\xi\rangle.$$

We graphically represent ev_H by drawing a cap.

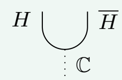


There is also a distinguished *coevaluation* map $\mathrm{coev}_H : \mathbb{C} \to H \otimes \overline{H}$ given by

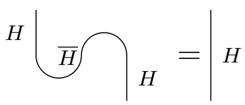
$$1_{\mathbb{C}} \longmapsto \sum |e_i\rangle \otimes \langle e_i|$$

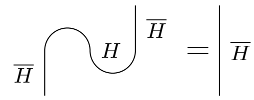
where $\{e_i\}$ is a choice of ONB. This sum is independent of the choice by Exercise 1.5.15. We graphically represent coev_H by drawing a cup.

$\mathrm{coev}_H =$ 

$=$ 

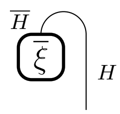
Exercise 1.6.10. Prove that ev_H and coev_H satisfy the *zig-zag/snake axioms*:



$= \mathrm{id}_H$ and 

$= \mathrm{id}_{\overline{H}}$.

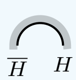
Exercise 1.6.11. Recall from Exercise 1.4.5 that we may view elements $\overline{\xi} \in \overline{H}$ as maps $\mathbb{C} \to \overline{H}$. Prove that

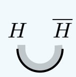


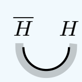
$= \langle \xi| : H \to \mathbb{C}.$

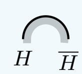
Notation 1.6.12 — Sometimes we use evaluations and coevaluations together with their adjoints in the same mathematical expression. When doing so, in order to easily differentiate between ev_H and $\mathrm{coev}_H^\dagger$ in the graphical calculus, we introduce a *framing* for strands corresponding to H and $H^\vee \cong \overline{H}$. We denote this framing by a lighter shaded thick line to the right side of the strand for H and the left side of $\overline{H}$ as follows.

We can then represent $\mathrm{ev}_H, \mathrm{coev}_H, \mathrm{ev}_H^\dagger, \mathrm{coev}_H^\dagger$ unambiguously as



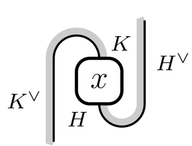
$= \mathrm{ev}_H$ 

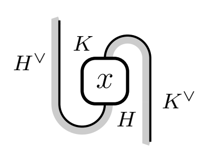
$= \mathrm{coev}_H$ 

$= \mathrm{ev}_H^\dagger$ 

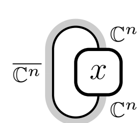
$= \mathrm{coev}_H^\dagger$.

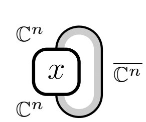
Exercise 1.6.13. Suppose $x : H \to K$. Prove that the transpose $x^\vee : K^\vee \to H^\vee$ is given by either rotation



$= x^\vee =$ 

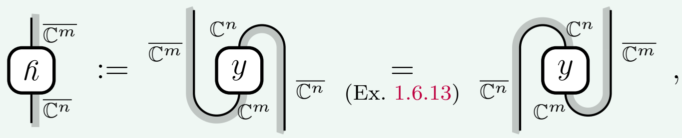
.

Exercise 1.6.14. For $x \in M_n(\mathbb{C})$, prove that 

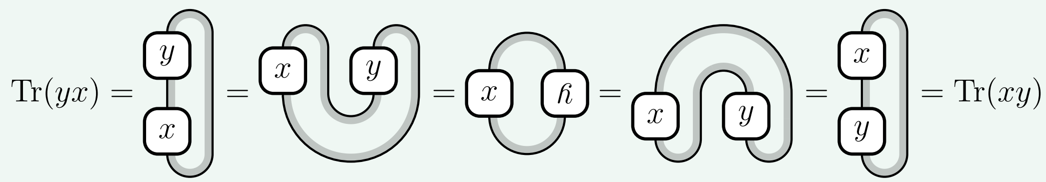
$= \mathrm{Tr}(x) =$ 

.

Example 1.6.15 — Using Exercises 1.6.13 and 1.6.14, we see that (1.6.1) is a well-typed proof. Indeed, for $y \in M_{n\times m}(\mathbb{C})$, setting

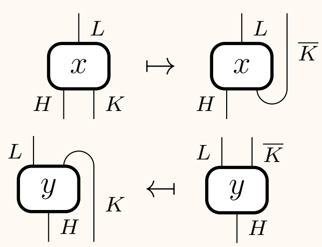


we can make the equation (1.6.1) more verbose.



Proposition 1.6.16 — The *Frobenius reciprocity* isomorphisms

$$B(H \otimes K \to L) \cong B(H \to L \otimes \overline{K})$$



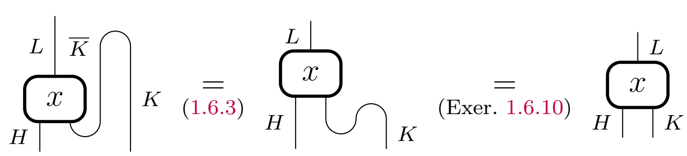
give a canonical isomorphism of vector spaces. Moreover, these maps are *natural* in H, L, i.e., for all $x : H_1 \to H_2$ and $y : L_1 \to L_2$ the following diagram commutes.

$$\begin{tikzcd}
B(H_2 \otimes K \to L_1) \arrow[d, leftrightarrow, "\cong"] & B(H_1 \otimes K \to L_1) \arrow[l, "-\circ(x\otimes \mathrm{id}_K)"'] \arrow[r, "y\circ -"] \arrow[d, leftrightarrow, "\cong"] & B(H_1 \otimes K \to L_2) \arrow[d, leftrightarrow, "\cong"] \\
B(H_2 \to L_1 \otimes \overline{K}) & B(H_1 \to L_1 \otimes \overline{K}) \arrow[l, "-\circ x"'] \arrow[r, "(y\otimes \mathrm{id}_{\overline{K}})\circ -"] & B(H_1 \to L_2 \otimes \overline{K})
\end{tikzcd}$$

Similarly, we have canonical Forbenius reciprocity unitary natural isomorphisms

$$B(H \to K \otimes L) \cong B(\overline{K} \otimes H \to L).$$

Proof. That these maps are mutually inverse follows by the exchange relation (1.6.3) and Exercise 1.6.10.



We leave the other composite to the reader. As only the K-strand wiggles, naturality follows immediately. $\square$

Corollary 1.6.17 — There is an canonical isomorphism of *vector spaces* $K \otimes \overline{H} \cong B(H \to K)$ given by $|\eta\rangle \otimes \langle\xi| \mapsto |\eta\rangle\langle\xi|$.

Proof. Apply Proposition 1.6.16 to

$$H \otimes \overline{K} \cong B(\mathbb{C} \to H \otimes \overline{K})$$

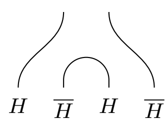
as vector spaces as in Exercise 1.4.5. $\square$

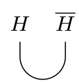
Exercise 1.6.18. Prove that the canonical isomorphism in Corollary 1.6.17 above can be promoted to a unitary isomorphism of Hilbert spaces if $B(H \to K)$ is equipped with the inner product

$$\langle x|y\rangle = \mathrm{Tr}_{B(H)}(x^\dagger y) = \mathrm{Tr}_{B(K)}(yx^\dagger).$$

Exercise 1.6.19. Prove that under the canonical unitary isomorphism $H \otimes \overline{H} \cong B(H)$ from Corollary 1.6.17 and Exercise 1.6.18, $\mathrm{coev}_H(1_{\mathbb{C}}) \mapsto \mathrm{id}_H$.

Exercise 1.6.20. Prove that under the unitary isomorphism $B(H) \cong H \otimes \overline{H}$ from Corollary 1.6.17 and Exercise 1.6.18, the multiplication and unit are given by

$m_{H\otimes\overline{H}} =$ 

$1_{H\otimes\overline{H}} =$ 

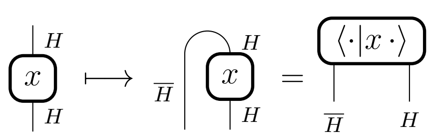
.

Write out $m_{H\otimes\overline{H}}$ and $1_{H\otimes\overline{H}}$ in bra-ket notation.

Corollary 1.6.21 (Operator-form correspondence) — The map $x \mapsto \langle\cdot|x\cdot\rangle$ is a bijective correspondence between operators $B(H)$ and sesquilinear forms such that:

- x is self-adjoint if and only if $\langle\cdot|x\cdot\rangle$ is self-adjoint,
- x is positive if and only if $\langle\cdot|x\cdot\rangle$ is positive,
- x is invertible if and only if $\langle\cdot|x\cdot\rangle$ is non-degenerate.

Proof. The second Frobenius reciprocity unitary 1.6.16 for $L = \mathbb{C}$ and $K = \overline{H}$ is given by



.

We leave the verification of the correspondence of properties as an exercise. $\square$

Exercise 1.6.22. Most of the facts above discussed for Hilbert spaces can be adapted to vector spaces using $V^\vee$ in place of $\overline{V}$. Go through this section again and rework it after forgetting the inner products.

1.7 Spectral theory

Definition 1.7.1 — Suppose $x \in B(H \to K)$ and $y : K \to H$. If $yx = 1_H$, then y is called a *left inverse* for x or a *retract*. If $xy = 1_K$, then y is called a *right inverse* for x or a *section*. If y is both a left and a right inverse, then y is called an *inverse* for x. We say x is *invertible* if it admits an inverse.

Exercise 1.7.2. Prove that if $x : H \to K$ has a left inverse y such that $yx = 1_H$ and a right inverse z such that $xz = \mathrm{id}_K$, then $y = z$. Deduce that inverses are unique when they exist.

By the above exercise, we denote the inverse of $x \in B(H)$ by x^{-1} when it exists.

Exercise 1.7.3. Suppose $x \in B(H)$.

(1) Show that if $\|1 - x\| < 1$, then x is invertible.
Hint: Prove that $\sum(1-x)^n$ *converges to* x^{-1} *in operator norm.*

(2) Deduce that if $x \in B(H)$ is invertible and $\|x - y\| < \|x^{-1}\|^{-1}$, then y is invertible.

(3) Deduce that the set of invertible operators $B(H)^\times$ is open and the map $x \mapsto x^{-1}$ is continuous on $B(H)^\times$.

Definition 1.7.4 — The *spectrum* of an operator $x \in B(H)$ is

$$\operatorname{spec}(x) := \{\lambda \in \mathbb{C} \mid \lambda - x \text{ is not invertible}\}.$$

This set is the same as the set of eigenvalues of x after identifying $B(H) \cong M_n(\mathbb{C})$, which is also the set of roots of the characteristic polynomial $\chi_x(\lambda) = \det(\lambda - x)$. Recall that $\operatorname{spec}(x) \neq \emptyset$ by the Fundamental Theorem of Algebra (every complex polynomial has a root).

Remark 1.7.5. Since every $\lambda \in \operatorname{spec}(x)$ is an eigenvalue of x, say with unit eigenvector ξ, observe that

$$|\lambda| = \|\lambda\xi\| = \|x\xi\| \leq \|x\|.$$

Exercise 1.7.6. Suppose (λ_n) is a bounded sequence in $\mathbb{C}$, and $F \subset \mathbb{C}$ is a finite set such that any convergent subsequence (λ_{n_k}) converges to a point in F. Prove that for any open neighborhood U of F, eventually $\lambda_n \in U$. That is, for every $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $n \geq N$ implies every λ_n is at most ε away from some $\lambda \in F$.

Proposition 1.7.7 (Continuity of spectrum) — If $x_n \to x$ in $B(H)$ in the operator norm, then for every open neighborhood U of $\operatorname{spec}(x)$, eventually $\operatorname{spec}(x_n) \subset U$.

Proof. Let U be an open neighborhood of $\operatorname{spec}(x)$ and $\mu \in U^c$. Suppose $\mu_n \to \mu$. Then $x_n - \mu_n \to x - \mu \in B(H)^\times$, which is open. Thus eventually $x_n - \mu_n \in B(H)^\times$, so eventually $\mu_n \notin \operatorname{spec}(x_n)$.

To see that this fact implies the result, observe that since $x_n \to x$, $\bigcup_n \operatorname{spec}(x_n)$ is a bounded set. Thus if (λ_n) is some sequence with $\lambda_n \in \operatorname{spec}(x_n)$, then for any convergent subsequence (λ_{n_k}), the limit of (λ_{n_k}) must be an element of $\operatorname{spec}(x)$. The result now follows by Exercise 1.7.6. □

Exercise 1.7.8. Suppose $x \in M_n(\mathbb{C})$ is normal. Prove that if $\lambda \in \operatorname{spec}(x)$ with corresponding eigenvector $\eta \in H$, then $\overline{\lambda} \in \operatorname{spec}(x^\dagger)$ with corresponding eigenvector η.
Hint: Use Exercise 1.4.20.

Theorem 1.7.9 (Spectral) — The following are equivalent $x \in M_n(\mathbb{C})$.

(1) There is an ONB of $\mathbb{C}^n$ consisting of eigenvectors for x.

(2) There is a unitary $u \in M_n(\mathbb{C})$ such that $u^\dagger x u$ is diagonal.

(3) x is normal.

Proof.

(1) ⇒ (2) : Let $\{e_j\}$ be such an ONB of eigenvectors for x, and set u to be the change of basis matrix

$$u := \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix}.$$

The eigenvalue equation implies $xu = ud$ where

$$d := \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

is the diagonal matrix whose entries are the corresponding eigenvalues of x. Then u is unitary as its columns are orthonormal (see Example 1.3.15), so $u^\dagger x u = d$.

(2) ⇒ (3) : When $d = u^\dagger x u$ is diagonal,

$$x^\dagger x = ud^\dagger u^\dagger u d u^\dagger = ud^\dagger d u^\dagger = udd^\dagger u^\dagger = udu^\dagger u d^\dagger u^\dagger = xx^\dagger.$$

(3) ⇒ (1) : Suppose x is normal and let $\lambda \in \operatorname{spec}(x)$ with eigenvector $\eta \in \mathbb{C}^n$. By Exercise 1.7.8, η is also an eigenvector of $x^\dagger$ with eigenvalue $\overline{\lambda} \in \operatorname{spec}(x^\dagger)$. Hence $\mathbb{C}\eta$ is invariant for x and $x^\dagger$. By Corollary 1.4.26, $xp = px$ where $p = |\eta\rangle\langle\eta|$. By Remark 1.4.27, x is diagonal with respect to the direct sum decomposition $\mathbb{C}^n = \operatorname{im}(p) \oplus \operatorname{im}(1-p)$. We now replace x by $(1-p)x = x(1-p)$ (which is again normal) acting on $\operatorname{im}(1-p) \subset \mathbb{C}^n$ (which has dimension $n-1$) and iterate the above procedure to obtain the desired ONB of eigenvectors. □

Exercise 1.7.10 (Real Spectral Theorem)**.** Prove that the following are equivalent for $x \in M_n(\mathbb{R})$

(1) There is an ONB of $\mathbb{R}^n$ consisting of eigenvectors for x.

(2) There is an *orthogonal* matrix $u \in M_n(\mathbb{R})$ satisfying $u^T u = 1 = uu^T$ such that $u^T xu$ is diagonal.

(3) x is *symmetric*, i.e., $x = x^T$.

Definition 1.7.11 (Functional calculus) — Suppose $x \in M_n(\mathbb{C})$ is normal. For $\lambda \in \operatorname{spec}(x)$ let $E_\lambda \subset \mathbb{C}^n$ denote the corresponding eigenspace, and let $p_\lambda \in M_n(\mathbb{C})$ be the orthogonal projection onto E_λ. We call the p_λ the *spectral projections* of x, and we note that they are mutually orthogonal ($p_\lambda p_\mu = 0$ for $\lambda \neq \mu$ in $\operatorname{spec}(x)$) and sum to 1.

Note that

$$x = \sum_{\lambda \in \operatorname{spec}(x)} \lambda p_\lambda \qquad \text{and} \qquad x^\dagger = \sum_{\lambda \in \operatorname{spec}(x)} \overline{\lambda} p_\lambda$$

as both operators agree on an orthonormal basis of $\mathbb{C}^n$, namely the orthonormal basis consisting of eigenvectors for x from the Spectral Theorem 1.7.9. For $f : \operatorname{spec}(x) \to \mathbb{C}$, we define

$$f(x) := \sum_{\lambda \in \operatorname{spec}(x)} f(\lambda) p_\lambda \in M_n(\mathbb{C}).$$

Observe that $\operatorname{spec}(f(x)) = f(\operatorname{spec}(x))$, as $f(x)$ is a diagonal operator with respect to the projections p_λ.

Exercise 1.7.12. For an operator $x \in M_n(\mathbb{C})$, we define its *spectral radius* to be

$$\rho(x) := \max \{|\lambda| \mid \lambda \in \operatorname{spec}(x)\}.$$

For x normal, show that $\|x\| = \rho(x)$.

At this point, we have not yet discussed the notion of a complex $*$-algebra; we will do so in Definition **??** below. However, the next theorem in best stated in these terms.

Theorem 1.7.13 (Gelfand) — Suppose $x \in M_n(\mathbb{C})$ is normal, and let $C(\operatorname{spec}(x))$ denote the unital $*$-algebra of $\mathbb{C}$-valued functions on $\operatorname{spec}(x)$. The map

$$C(\operatorname{spec}(x)) \ni f \mapsto f(x) \in M_n(\mathbb{C})$$

is an isometric unital $*$-algebra homomorphism onto the unital $*$-algebra generated by x. That is, for all $f, g \in C(\operatorname{spec}(x))$,

$$(f+g)(x) = f(x) + g(x), \qquad (fg)(x) = f(x)g(x), \qquad \text{and} \qquad \overline{f}(x) = f(x)^\dagger,$$

and $\|f(x)\| = \|f\|_{C(\mathrm{spec}(x))} := \max\{|f(\lambda)| \mid \lambda \in \mathrm{spec}(x)\}$.

Proof. It is straightforward to verify that $f \mapsto f(x)$ is a unital $*$-algebra map by checking the action of $f(x)$ on the ONB of eigenvectors of x from the Spectral Theorem 1.7.9. Injectivity follows as $f \neq g$ on $\mathrm{spec}(x)$ implies that $f(\lambda)p_\lambda \neq g(\lambda)p_\lambda$ for some $\lambda \in \mathrm{spec}(x)$. Since the image contains 1, x, and $x^\dagger$ by construction, it is onto the unital $*$-algebra generated by x. The last claim follows by Exercise 1.7.12 since $f(x)$ is again normal. $\square$

Exercise 1.7.14. Use the functional calculus to prove that every positive $x \in M_n(\mathbb{C})$ has a unique positive square root. That is, if $x \geq 0$, there is a unique positive operator $\sqrt{x} \in M_n(\mathbb{C})$ such that $\sqrt{x}^2 = x$.

Proposition 1.7.15 (Continuity of functional calculus) — Suppose $x_n \to x$ is a convergent sequence of normal operators in operator norm and $f : U \to \mathbb{C}$ is a continuous function on some open neighborhood U of $\mathrm{spec}(x)$. Then $f(x_n)$ is eventually well-defined by Proposition 1.7.7, and $f(x_n) \to f(x)$.

Proof. Clearly the result holds for polynomials $p \in \mathbb{C}[z, \overline{z}]$. Now by the classical Stone-Weierstrass theorem, we may uniformly approximate f by a polynomial $p \in \mathbb{C}[z, \overline{z}]$. Thus

$$\|f(x) - f(x_n)\| \leq \underbrace{\|f(x) - p(x)\|}_{=\|f-p\|_{C(\mathrm{spec}(x))}} + \underbrace{\|p(x) - p(x_n)\|}_{\xrightarrow{n\to\infty} 0} + \underbrace{\|p(x_n) - f(x_n)\|}_{=\|f-p\|_{C(\mathrm{spec}(x_n))}}$$

Since $\bigcup \mathrm{spec}(x_n)$ is bounded, we can make $\|f-p\|_{C(\mathrm{spec}(x))}, \|f-p\|_{C(\mathrm{spec}(x_n))}$ uniformly small over all n, and the result follows. $\square$

Exercise 1.7.16. In this exercise, we prove that the set $GL_+(n)$ of positive invertible $n \times n$ matrices is *contractible*.

(1) Prove that any real vector space V is contractible via the *strong deformation retraction* $[0,1] \times V \to V$ given by $(t, \xi) \mapsto t\xi$. Deduce that $M_n(\mathbb{C})_{\mathrm{sa}}$ is contractible.

(2) Prove that $\exp : M_n(\mathbb{C})_{\mathrm{sa}} \to GL_+(n)$ is a homeomorphism with inverse $\log : GL_+(n) \to M_n(\mathbb{C})_{\mathrm{sa}}$. Deduce that $GL_+(n)$ is contractible.

Proposition 1.7.17 — Suppose $x, y \in M_n(\mathbb{C})$ with x normal and $xy = yx$. Then $f(x)y = yf(x)$ for every $f \in C(\mathrm{spec}(x))$.

Proof. Since $\mathrm{spec}(x)$ is a finite set, there is a polynomial p such that $p = f$ on $\mathrm{spec}(x)$. Since $x^n y = yx^n$ for every n, $p(x)y = yp(x)$, and the result follows. $\square$

Exercise 1.7.18. Let $x \in M_n(\mathbb{C})$.

(1) Show that x can be written uniquely as $\mathrm{Re}(x) + i\,\mathrm{Im}(x)$ where both $\mathrm{Re}(x), \mathrm{Im}(x)$ are self-adjoint.

(2) Show that if x is self-adjoint, then x can be written uniquely as $x = x_+ - x_-$ where x_+, x_- are both positive and $x_+ x_- = 0$.

(3) Show that if x is self-adjoint, then $x \leq \|x\|$, i.e., $\|x\| - x \geq 0$.

(4) Show that if x is self-adjoint, then x can be written as a linear combination of two unitaries in $M_n(\mathbb{C})$.
Hint: if $\|x\| \leq 1$, consider $u := x + i\sqrt{1-x^2}$.

Proposition 1.7.19 — If $x \geq 0$, then any unit vector $\eta \in \mathbb{C}^n$ which maximizes

$$x \longmapsto \langle \eta | x\eta \rangle$$

is an eigenvector corresponding to $\rho(x) = \|x\|$ (see Exercise 1.7.12).

Proof. By Exercise 1.7.18, $\|x\| - x \geq 0$. Thus

$$\langle \eta | x\eta \rangle = \|x\| \qquad \Longleftrightarrow \qquad 0 = \langle \eta | (\|x\| - x)\eta \rangle = \|(\|x\| - x)^{1/2}\eta\|$$

which implies $(\|x\| - x)^{1/2}\eta = 0$. Multiplying by $(\|x\| - x)^{1/2}$, we see that $x\eta = \|x\|\eta$. □

Proposition 1.7.20 — The following are equivalent for $x \in M_n(\mathbb{C})$.

(1) $x \geq 0$.

(2) x is normal and all eigenvalues of x are non-negative.

(3) There exists $y \in M_n(\mathbb{C})$ such that $y^\dagger y = x$.

(4) There exists $y \in M_{n\times k}(\mathbb{C})$ for some $k \in \mathbb{N}$ such that $y^\dagger y = x$.

Proof.

$\underline{(1) \Rightarrow (2):}$ Positive implies self-adjoint by Corollary 1.4.19, and self-adjoint clearly implies normal. If η is an eigenvector of x with eigenvalue λ, $0 \leq \langle \eta | x\eta = \lambda \langle \eta | \eta \rangle$, so $\lambda \geq 0$.

$\underline{(2) \Rightarrow (3):}$ Use the functional calculus to define $\sqrt{x} \in M_n(\mathbb{C})$ as in Exercise 1.7.14 above. Observe $\sqrt{x}$ is self-adjoint and satisfies $\sqrt{x}^2 = x$.

$\underline{(3) \Rightarrow (4):}$ Trivial.

$\underline{(4) \Rightarrow (1):}$ Observe that for all $\eta \in \mathbb{C}^n$, $\langle \eta | x\eta \rangle_{\mathbb{C}^n} = \langle \eta | y^\dagger y\eta \rangle_{\mathbb{C}^n} = \langle y\eta | y\eta \rangle_{\mathbb{C}^k} \geq 0$. □

Definition 1.7.21 — For an operator $x \in M_n(\mathbb{C})$ its *support projection* is

$$\operatorname{supp}(x) := 1 - p_{\ker(x)}$$

where $p_{\ker(x)}$ is the orthogonal projection onto $\ker(x)$. Observe that $x = x\,\mathrm{supp}(x)$.

Remark 1.7.22. When x is normal, $x = x\,\mathrm{supp}(x) = \mathrm{supp}(x)x$, and $\mathrm{supp}(x)$ is the sum of all spectral projections of x except for p_0 if $0 \in \mathrm{spec}(x)$. Lemma 1.4.17 allows us to replace x for the normal operator $x^\dagger x$ in general. Thus $\mathrm{supp}(x)$ is well-defined independent of the action of $M_n(\mathbb{C})$ on $\mathbb{C}^n$.

Definition 1.7.23 (Polar decomposition) — Suppose $x \in M_{m\times n}(\mathbb{C})$. Using functional calculus, we define $|x| := \sqrt{x^\dagger x}$. The map $u : |x|\xi \mapsto x\xi$ on $\mathrm{supp}(x)\mathbb{C}^n$ and $u = 0$ on $(1 - \mathrm{supp}(x))\mathbb{C}^n$ is an isometric linear operator and thus well-defined:

$$\| |x|\xi \|^2 = \langle\, |x|\xi \,|\, |x|\xi \,\rangle = \langle\, \xi \,|\, |x|^2\xi \,\rangle = \langle \xi | x^\dagger x\xi\rangle = \langle x\xi | x\xi\rangle = \|x\xi\|^2.$$

Hence we may write $x = u|x|$ where u is a partial isometry and $|x| \geq 0$; this is called the *polar decomposition* of x.

Remark 1.7.24. When $x \in M_n(\mathbb{C})$, the partial isometry u constructed above commutes with all unitaries $v \in M_n(\mathbb{C})$ which commute with x and $x^\dagger$. Indeed, such a v commutes with $x^\dagger x$, and thus with $|x|$ and $\mathrm{supp}(|x|)$ by Proposition 1.7.17. This means $uv = 0 = vu$ on $(1 - \mathrm{supp}(x))\mathbb{C}^n$ and on $\mathrm{supp}(x)\mathbb{C}^n$,

$$vuv^*|x|\xi = vu|x|v^*\xi = vxv^*\xi = x\xi.$$

Thus $vuv^* = u$, so $vu = uv$.

Exercise 1.7.25. In this exercise, we will prove the uniqueness of the polar decomposition.

(1) Prove that u is the unique partial isometry such that $x = u|x|$ and $\ker(u) = \ker(x)$.

(2) Deduce that u is the unique partial isometry such that with $x = u|x|$ and $u^\dagger u = \mathrm{supp}(|x|)$. In this sense, the polar decomposition is independent of the action of $M_{m\times n}(\mathbb{C})$ on $\mathbb{C}^n$.

(3) Deduce that if $x = vy$ where $y \geq 0$ with $\mathrm{supp}(y) = \mathrm{supp}(|x|)$ and v is a partial isometry with $v^\dagger v = \mathrm{supp}(|x|)$, then $y = |x|$ and $v = u$.

Lemma 1.7.26 — Suppose $x \in M_{m\times n}(\mathbb{C})$, and let $x = u|x|$ be its polar decomposition.

(1) $u^\dagger u = \mathrm{supp}(x)$ and $uu^\dagger = \mathrm{supp}(x^\dagger)$, and

(2) $u^\dagger x = |x|$ and $x = |x^\dagger|u$, and

(3) the polar decomposition of $x^\dagger$ is given by $u^\dagger|x^\dagger|$.

Proof.

(1): First, since $\ker(u) = \ker(x)$, $u^\dagger u = 1 - p_{\ker(x)} = \operatorname{supp}(x)$.

Second, since $x^\dagger = |x|u^\dagger$, $\ker(u^\dagger) \subseteq \ker(x^\dagger)$. If $\eta \in \ker(x^\dagger)$, then $0 = x^\dagger\eta = |x|u^\dagger\eta$, so $u^\dagger\eta \in \ker(|x|) = \ker(u)$. Hence $uu^\dagger\eta = 0$, so $\eta \in \ker(uu^\dagger) = \ker(u^\dagger)$ by Lemma 1.4.17. Thus $\ker(x^\dagger) = \ker(u^\dagger)$, so $uu^\dagger = 1 - p_{\ker(x^\dagger)} = \operatorname{supp}(x^\dagger)$.

(2): Since $\ker(x) = \ker(|x|)$, $\operatorname{supp}(|x|) = u^\dagger u$ by (1). Thus $u^\dagger x = u^\dagger u|x| = \operatorname{supp}(|x|)|x| = |x|$ by Remark 1.7.22.

Since $uu^\dagger|x^\dagger| = |x^\dagger|$ and $|x|u^\dagger u = |x|$,

$$(u^\dagger|x^\dagger|u)^2 = u^\dagger|x^\dagger|uu^\dagger|x^\dagger|u = u^\dagger|x^\dagger|^2u = u^\dagger xx^\dagger u = |x|^2 = x^\dagger x.$$

Hence $u^\dagger|x^\dagger|u = |x|$ by uniqueness of the positive square root (Exercise 1.7.14). Hence

$$x = u|x| = uu^\dagger|x^\dagger|u = \operatorname{supp}(|x^\dagger|)|x^\dagger|u = |x^\dagger|u.$$

(3): Taking $\dagger$ in the second equation in (2) gives $x^\dagger = u^\dagger|x^\dagger|$. Since we showed $\ker(u^\dagger) = \ker(x^\dagger)$ in (1), it is indeed the polar decomposition. $\square$

Corollary 1.7.27 — For $x \in M_{m\times n}(\mathbb{C})$, the following are equivalent.

(1) x has a left inverse.

(2) $x^\dagger x$ is invertible.

(3) In the polar decomposition $x = u|x|$, u is an isometry.

Dually, x has a right inverse if and only if $xx^\dagger$ is invertible if and only if u is a coisometry.

Proof. Since $\ker(x) = \ker(x^\dagger x)$, x has a left inverse if and only if $\ker(x^\dagger x) = \ker(x) = 0$ if and only if $x^\dagger x$ is invertible by the Rank-Nullity Theorem 1.1.17. Moreover, $\ker(x) = 0$ if and only if $u^\dagger u = 1 - p_{\ker(x)} = 1$.

The dual statement for (1) $\Leftrightarrow$ (2) follows formally by considering $x^\dagger$. The dual statement for (2) $\Leftrightarrow$ (3) follows as $\ker(x^\dagger) = 0$ if and only if $uu^\dagger = 1$ by part (3) of Lemma 1.7.26. $\square$