Mathematical preliminaries

1 Hilbert spaces and inner products

To every quantum (-mechanical) system one can associate a Hilbert space $\mathcal{H}$. In these lectures, we will only consider quantum systems whose associated Hilbert space is finite dimensional: $\dim \mathcal{H} = d < \infty$.

A Hilbert space is a real or complex inner product space that is also a complete1 metric space with regard to the distance function induced by the inner product.

- That $\mathcal{H}$ is a complex inner product space means that it is a complex vector space on which there is an inner product (or scalar product)

$$
(x, y) \in \mathbb{C}, \quad \forall x, y \in \mathcal{H},
$$

satisfying the following properties

1. $$(y, x) = \overline{(x, y)}$$ (where the bar denotes the complex conjugate of a complex number)

2. Linear in the second argument

$$
(x, ay_1 + by_2) = a(x, y_1) + b(x, y_2)
$$

Note that this property along with 1 implies antilinearity in the first argument:

$$
(ax_1 + bx_2, y) = \overline{a(y, x_1)} + \overline{b(y, x_2)} = \overline{a(x_1, y)} + \overline{b(x_2, y)}.
$$

3. Positive definite: $(x, x) \geq 0$ with ‘$=$’ only when $x = 0$ (the zero vector in $\mathcal{H}$).

Note: a real inner product space is defined similarity except that $\mathcal{H}$ is a real vector space and the inner product takes real values. Such an inner product is linear in each argument.

- The norm of $x \in \mathcal{H}$ is a real-valued function given by

$$
\|x\| = \sqrt{(x, x)}.
$$

1i.e. all Cauchy sequences converge

- The distance between two vectors in the Hilbert space is given by

$$
d(x, y) = \|x - y\|. \tag{3}
$$

d is called a metric, and

a) is symmetric

b) has $d(x, y) = 0$ iff $x = 0$

c) satisfies the triangle inequality

$$
d(x, z) \leq d(x, y) + d(y, z).
$$

Cauchy-Schwarz inequality: for all $x, y \in \mathcal{H},$

$$
|\langle x, y \rangle| \leq \sqrt{(x, x)} \sqrt{(y, y)}.
$$

We denote the Hilbert space associated to a quantum system $A$ as $\mathcal{H}_A$ and its dimension $\dim \mathcal{H}_A$ as $d_A$.

2 Linear maps / operators on Hilbert spaces

- Consider a linear operator $A : \mathcal{H} \rightarrow \mathcal{H}'$. The set of such linear operators is denoted $\mathcal{B}(\mathcal{H}, \mathcal{H}')$. Then $A \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$, and $A$ is a homomorphism in the set of homomorphisms from $\mathcal{H}$ to $\mathcal{H}'$, $\text{Hom}(\mathcal{H}, \mathcal{H}')$.

- A homomorphism (w.r.t. the inner product): if $A : \mathcal{H} \rightarrow \mathcal{H}'$, $A \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$, then its adjoint

$$
A^\dagger : \mathcal{H}' \rightarrow \mathcal{H}, \quad A^\dagger \in \mathcal{B}(\mathcal{H}', \mathcal{H}) \tag{4}
$$

is the unique operator defined by

$$
\langle v, Av' \rangle = \langle A^\dagger v, v' \rangle. \tag{5}
$$

In particular, $(A^\dagger)^\dagger = A$.

- Note that $\mathcal{B}(\mathcal{H})$ itself can be elevated to a real Hilbert space by equipping it with the Hilbert-Schmidt inner product $\langle A, B \rangle = \text{Tr}[A^\dagger B]$. 

\[ \text{Part III, Lent 2019} \]
\[ \text{University of Cambridge} \]
\[ \text{Quantum Information Theory} \]
Matrix representation: given a (orthonormal\textsuperscript{2}) basis \(\{v_i\}_{i=1}^{d}\) of \(\mathcal{H}\), an operator \(A \in \mathcal{B}(\mathcal{H})\) can be represented by a matrix with elements

\[ A_{ij} = \langle v_i, Av_j \rangle \] (6)

where \(i\) is the row index and \(j\) is the column index. The matrix representation isn’t unique; it depends on the choice of basis. If \(\mathcal{H} = \mathbb{C}^d\), then \(\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathbb{C}^d) = \mathcal{M}_d\), the set of \(d \times d\) complex matrices. For \(d = 2\) and

\[
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \quad A^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}
\] (7)

is the conjugate transpose. The identity operator \(I\) is the identity matrix in any basis:

\[
I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\] (8)

Properties of \(A \in \mathcal{B}(\mathcal{H})\):

- If \(A\) is self-adjoint / Hermitian if \(A = A^\dagger\).
- \(A\) is normal if \(AA^\dagger = A^\dagger A\).
- \(A\) is positive semi-definite if \((v, Av) \geq 0\) for all \(v \in \mathcal{H}\). This is denoted \(A \succeq 0\). Note that in particular, for \(A\) to be positive semi-definite, \((v, Av)\) must be real for all \(v \in \mathcal{H}\). This implies \(A = A^\dagger\).
- \(A\) is a projector if \(A \geq 0\) and \(A = A^2\).
- \(\mathcal{B}_n(\mathcal{H}) = \{ A \in \mathcal{B}(\mathcal{H}) : A = A^\dagger \}\).

Dirac’s bra-ket notation: We write a vector \(v \equiv |v\rangle \in \mathcal{H}\) which is represented by a column vector (\(\mathcal{H} \cong \mathbb{C}^d\)),

\[
|v\rangle = \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix}.
\] (9)

Its dual is the conjugate transpose

\[
\langle v | = (a_1^*, \ldots, a_d^*). \] (10)

\textsuperscript{2}We will only consider orthonormal bases.

Another vector \(v' = |v'\rangle = \begin{pmatrix} b_1 \\ b_d \end{pmatrix}\).

Then

\[
(v, v') = (|v\rangle |v'\rangle) = (a_1^*, \ldots, a_d^*) \begin{pmatrix} b_1 \\ \vdots \\ b_d \end{pmatrix} = \sum_{i=1}^{d} a_i^* b_i.
\] (11)

Identity operator

\[
1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\] (12)

in any basis. So for all bases \(\{|e_i\}_{i=1}^{d}\) and \(\{|f_i\}_{i=1}^{d}\),

\[
1 = \sum_{i=1}^{d} |e_i\rangle \langle e_i| = \sum_{i=1}^{d} |f_i\rangle \langle f_i|.
\] (13)

This is the outer product representation. Then using this and writing \(A = 1 A 1\), you can see the non-uniqueness of the matrix representation.

2.1 For those interested: duality between bras and kets (nonexaminable)

Given a vector \(v \in \mathcal{H}\), we can consider \(|v\rangle\) as a map from \(\mathbb{C} \to \mathcal{H}\), i.e. \(|v\rangle \in \text{Hom}(\mathbb{C}, \mathcal{H})\), by regarding \(|v\rangle\) as mapping a complex number \(z \in \mathbb{C}\) to the vector \(z v \in \mathcal{H}\). Then we can consider the adjoint map \(|v\rangle^\dagger \in \text{Hom}(\mathcal{H}, \mathbb{C})\). By definition, \(|v\rangle^\dagger\) is a linear map such that for all \(z \in \mathbb{C}\) and \(w \in \mathcal{H}\),

\[
(|v\rangle(z), w)_H = (z v, w)_H = (z, |v\rangle^\dagger(w))_\mathbb{C}
\] (14)

where we have used subscripts to distinguish between inner products on \(\mathcal{H}\) and inner products on \(\mathbb{C}\). Taking \(z = 1\), we see that we require \((v, w)_H = |v\rangle^\dagger(w)\) for all \(w \in \mathcal{H}\). That is, \(|v\rangle\) is the operator that acts on vectors in \(\mathcal{H}\) by taking the inner product with \(v\). Let us denote \(|v\rangle^\dagger = |v\rangle^\ast\).

Then \(
\langle v | w \rangle := \langle v | \circ | w \rangle \in \text{Hom}(\mathbb{C}, \mathbb{C})\) is a linear operator on \(\mathbb{C}\). How does it act? For \(z \in \mathbb{C}\),

\[
\langle v | w \rangle(z) = (v | zw) = z (v | w)_H.
\] (15)

That is, \(|v\rangle |w\rangle\) is the linear operator on \(\mathbb{C}\) that acts as scalar multiplication by \((v, w)_H\). We just identify this operator with the number \((v, w)_H\). Often we write \(\langle v | w \rangle\) for this quantity.
Clearly, for any $v \in \mathcal{H}$ we may define such a linear operator $\langle v |$ which acts by taking the inner product with $v$. On the other hand, the Riesz representation theorem (for Hilbert spaces) gives that any element of $\text{Hom}(\mathcal{H}, \mathbb{C})$ is of this form. This means that the set of bras is on an equal footing with the set of kets: it forms a complete inner product space as well, and in fact, is isomorphic to the set of kets (via the map $|v\rangle \to \langle v |$).

### 3 Infimum and Supremum

Consider a set of real numbers $S \subseteq \mathbb{R}$, which is nonempty. The set $S$ has an upper bound if there is a real number $u \in \mathbb{R}$ such that $s \leq u$ for every $s \in S$. Likewise, $S$ has a lower bound if there is a real number $l \in \mathbb{R}$ such that $l \leq s$ for all $s \in S$. The set $S$ is bounded above if it has an upper bound, and bounded below if it has a lower bound.

If $S$ is bounded above, then the supremum of $S$ is a real number $u \in \mathbb{R}$ such that

a) $u$ is an upper bound of $S$, and

b) $u$ is the least upper bound of $S$: if $v$ is any other upper bound of $S$, then $u \leq v$.

The supremum of $S$ is denoted $\sup S$ (i.e. $\sup S = u$ satisfying the two points above).

Similarly, if $S$ is bounded below, then the infimum of $S$ is a real number $l \in \mathbb{R}$ such that

a) $l$ is a lower bound of $S$, and

b) $l$ is the greatest lower bound of $S$: if $k$ is any other lower bound of $S$, then $l \geq k$.

The infimum of $S$ is denoted $\inf S$.

**Some facts about infimum and supremum:**

- If $S \subseteq \mathbb{R}$ is bounded above, then the supremum exists and is unique, and if it is bounded below, the infimum exists and is unique.

- The supremum or infimum does not need to be an element of the set. For example, $\inf(0, 1] = 0$ but $0 \not\in (0, 1]$. On the other hand, $\sup(0, 1] = 1 \in (0, 1]$.

See e.g. Chapter 2 of [1] for more examples and definitions.

### 4 Concave and convex functions

A set $C \subseteq \mathbb{R}^n$ is convex if the line segment between any two points in $C$ lies in $C$. That is, for all $x_1, x_2 \in C$ and any $t$ with $0 \leq t \leq 1$, we have

$$tx_1 + (1-t)x_2 \in C.$$ 

(16)

A function $f : D \to \mathbb{R}$ defined on some domain $D \subseteq \mathbb{R}^n$ is convex if $D$ is a convex set, and if for all $x, y \in D$, and any $t$ with $0 \leq t \leq 1$, we have

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$ 

(17)

By induction, we can then prove Jensen’s inequality: if $y_1, \ldots, y_n \in D$ and $p_1, \ldots, p_n \in (0, 1]$ such that $\sum_{i=1}^n p_i = 1$, then

$$f\left(\sum_{i=1}^n p_i x_i\right) \leq \sum_{i=1}^n p_i f(x_i)$$

(18)

with equality if and only if $x_1 = x_2 = \ldots = x_n$ or $f$ is linear. See Chapter 3 of [2] for more on convex functions.

### References
