Quantum Information Theory
Michaelmas term 2015-16

Will Matthews

October 16, 2015

1 Course information

- My email address is `wm266@cam.ac.uk`.
- I will be available in my office (Pavilion F 0.14) every Monday 13:00 - 15:00.
- The course webpage is located at `http://northala.net/qit/`. Lecture notes and examples sheets will be posted here. For the most up-to-date information regarding the course, please look here first.
- Felix Leditzky (`f.leditzky@statslab.cam.ac.uk`) will run the examples classes. The plan is to have four examples classes in total. Three will be this term, from 14:00 to 16:00 in MR14 on

  – Friday 30th Oct,

  – Friday 13th Nov,

  – Friday 27th Nov.

  The fourth will be early next term (Lent) at a date to be determined.
- Please fill out the attendance list.
- I will add you to a mailing list which I will use to send out solutions for the examples class questions and for any important messages regarding the course (e.g. changes to examples class dates). If you want to unsubscribe at any time, just mail me.
- Exercises in the handouts are marked with ♣♣. These exercises will also appear on the examples sheets.

2 Background

2.1 Hilbert spaces

A complex Hilbert space $\mathcal{H}$ is a complex vector space equipped with an inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ and which is complete in the norm

$$\|v\| = \langle v, v \rangle^{1/2}$$

induced by the inner product. Recall that an inner product is

1. Linear in 2nd argument[1]: $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$, $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$.

2. Conjugate symmetric: $\langle x, y \rangle = \langle y, x \rangle^*$.

3. Positive definite: $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$.

Note that (1) and (2) imply conjugate linearity in the first argument: $\langle y+z, x \rangle = \langle y, x \rangle + \langle z, x \rangle$, $\langle \alpha y, x \rangle = \alpha^* \langle y, x \rangle$.

In this course we will only deal with Hilbert spaces of finite dimension, so the completeness condition is automatically satisfied. From now on we will nearly always write vectors in $\mathcal{H}$ as **'kets'** e.g. $|\psi\rangle$. If a Hilbert space is called $\mathcal{H}_\mathsf{Q}$, for some Q, then we will sometimes label vectors in $\mathcal{H}_\mathsf{Q}$ with the same subscript (e.g. $|\psi\rangle_\mathsf{Q}$) to indicate where they live, and write d_Q for the dimension $\dim(\mathcal{H}_\mathsf{Q})$ of $\mathcal{H}_\mathsf{Q}$.

2.1.1 Computational basis

We assume that each Hilbert space comes equipped with an orthonormal basis $\{|i\rangle : i = 0, \ldots, \dim(\mathcal{H}) - 1\}$ which we declare to be *real* vectors, i.e. $|i\rangle^* = |i\rangle$ for all $0 \leq i < \dim(\mathcal{H})$, and which we call the **computational basis**. If we write $|\psi\rangle \in \mathcal{H}$ as a column vector, this consists of the components of $|\psi\rangle$ in the computational basis (unless otherwise specified) e.g. if $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$ then $|\psi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}$. The orthonormality of the computational basis means that if $|\psi\rangle = \sum_{0 \leq j < \dim(\mathcal{H})} \alpha_j |j\rangle$ and $|\phi\rangle = \sum_{0 \leq j < \dim(\mathcal{H})} \beta_j |j\rangle$ then $\langle |\psi\rangle, |\phi\rangle \rangle = \sum_{0 \leq j < \dim(\mathcal{H})} \alpha_j^* \beta_j$.

2.2 Linear maps between Hilbert spaces

Given two spaces $\mathcal{H}_\mathsf{A}$ and $\mathcal{H}_\mathsf{B}$ we denote the complex vector space of linear maps from $\mathcal{H}_\mathsf{A}$ to $\mathcal{H}_\mathsf{B}$ by $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$. We call elements of $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A})$ *operators* on $\mathcal{H}_\mathsf{A}$ and use the abbreviation $\mathcal{L}(\mathcal{H}_\mathsf{A}) := \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A})$.

[1] This is the usual convention in physics texts, and matches the bra-ket notation.

2.2.1 Bras; Bra-kets

The **dual space** to $\mathcal{H}_\mathsf{A}$ is the vector space $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathbb{C})$ of complex linear functionals on $\mathcal{H}_\mathsf{A}$. The inner product on $\mathcal{H}_\mathsf{A}$ provides a natural way to associate to every $|\psi\rangle$ in $\mathcal{H}_\mathsf{A}$ a unique linear functional $|\psi\rangle^\dagger$, which is defined by $|\psi\rangle^\dagger : |\phi\rangle \mapsto \langle|\psi\rangle, |\phi\rangle\rangle$. We write $|\psi\rangle^\dagger$ as a **'bra'** $\langle\psi| := |\psi\rangle^\dagger$. Conversely, for every linear functional $f \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathbb{C})$ there is a unique vector $f^\dagger$ in $\mathcal{H}_\mathsf{A}$ such that $f(|\phi\rangle) = \langle f^\dagger, |\phi\rangle\rangle$. Therefore, $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathbb{C}) = \{\langle\psi| : |\psi\rangle \in \mathcal{H}_\mathsf{A}\}$ and $\langle\psi|^\dagger = |\psi\rangle$. We will abbreviate $\langle\psi||\phi\rangle$ to $\langle\psi|\phi\rangle$ (this is called a **'bra-ket'**), and will normally make use of the identity $\langle\psi|\phi\rangle = \langle|\psi\rangle, |\phi\rangle\rangle$ to express inner products.

2.2.2 Computational basis for linear maps

The computational basis for the dual space to $\mathcal{H}_\mathsf{A}$ is $\{\langle j| : 0 \le j < d_\mathsf{A}\}$ and we will write elements of the dual space $\langle\chi|_\mathsf{A} = \sum_{0\le j<d_\mathsf{A}} c_j\langle j|_\mathsf{A}$ as row vectors $\langle\chi|_\mathsf{A} = (c_0\ c_1\ \cdots\ c_{d-1})$. If $|\psi\rangle_\mathsf{A} = \sum_{0\le j<d_\mathsf{A}} \alpha_j|j\rangle$ then $\langle\psi|_\mathsf{A} = \sum_{0\le j<d_\mathsf{A}} \alpha_j^*\langle j|_\mathsf{A} = (\alpha_0^*\ \alpha_1^*\ \cdots\ \alpha_{d-1}^*)$, which is the conjugate transpose of the column vector associated to $|\psi\rangle_\mathsf{A}$ (we will denote conjugate transpose of a matrix with a †, also). The set of maps $\{|j\rangle_\mathsf{B}\langle k|_\mathsf{A} : 0 \le j < d_\mathsf{B}, 0 \le k < d_\mathsf{A}\}$ comprise the *computational basis* for $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$. If $X = \sum_{0\le j<d_\mathsf{B}}\sum_{0\le k<d_\mathsf{A}} X_{jk}|j\rangle_\mathsf{B}\langle k|_\mathsf{A}$ then we will write it as a $d_\mathsf{B} \times d_\mathsf{A}$ matrix with entries X_{jk}.

2.2.3 Hermitian adjoint

Given $X \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$, the **hermitian adjoint** of X is the unique operator $X^\dagger$ in $\mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{A})$ such that $\langle|\phi\rangle_\mathsf{B}, X|\psi\rangle_\mathsf{A}\rangle = \langle X^\dagger|\phi\rangle_\mathsf{B}, |\psi\rangle_\mathsf{A}\rangle$ for all $|\phi\rangle_\mathsf{B} \in \mathcal{H}_\mathsf{B}$ and $|\psi\rangle_\mathsf{A} \in \mathcal{H}_\mathsf{A}$. If $X = \sum_{j,k} X_{jk}|j\rangle_\mathsf{B}\langle k|_\mathsf{A}$ then $X^\dagger = \sum_{j,k} X_{jk}^*|k\rangle_\mathsf{A}\langle j|_\mathsf{B}$, and the matrix representation for $X^\dagger$ is the conjugate transpose of X. Note that the hermitian adjoint of a complex number (considered as an operator on a one-dimensional Hilbert space) is just its complex conjugate. We have the following identities: (1) $(X|\psi\rangle)^\dagger = \langle\psi|X^\dagger$, (2) $(\langle\psi|X)^\dagger = X^\dagger|\psi\rangle$, (3) $(XY)^\dagger = Y^\dagger X^\dagger$.

2.2.4 Hermitian operators

We say an operator $X \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$ is **hermitian** if $X^\dagger = X$. The set $\mathrm{Herm}(\mathcal{H}_\mathsf{Q})$ of hermitian operators on $\mathcal{H}_\mathsf{Q}$, is a *real* vector space of dimension $\dim(\mathsf{Q})^2$.

We call $E \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$ an **orthogonal projection operator** (or just **'projector'**) if it satisfies $E^\dagger E = E$ (equivalently, E is hermitian and $E^2 = E$).

Theorem 1 (Eigendecomposition)**.** Any hermitian operator $X \in \mathrm{Herm}(\mathcal{H}_\mathsf{Q})$ has a unique decomposition

$$X = \sum_{\lambda\in\mathrm{spec}(X)} \lambda\Pi_\lambda$$

where $\mathrm{spec}(X)$ is the set of eigenvalues of X and Π_λ is the orthogonal projector onto the eigenspace corresponding to eigenvalue λ. $\mathrm{spec}(X) \subset \mathbb{R}$ and

$$\Pi_\lambda\Pi_\mu = \begin{cases} \Pi_\lambda \text{ if } \lambda = \mu, \\ 0 \text{ if } \lambda \ne \mu. \end{cases}$$

Equivalently, any hermitian operator $X \in \mathrm{Herm}(\mathcal{H}_\mathsf{Q})$ can be written as $X = \sum_{0\le j<d_\mathsf{Q}} \lambda_j|a_j\rangle\langle a_j|$ where $|a_j\rangle$ is an eigenvector of X corresponding to eigenvalue λ_j. The eigenvalues are all real and the eigenvectors $\{|a_j\rangle : 0 \le j < d_\mathsf{Q}\}$ form an orthonormal basis for $\mathcal{H}_\mathsf{Q}$.

2.2.5 Support of an operator

The **support** of an operator $X \in \mathcal{L}(\mathcal{H})$ is the subspace $\mathrm{supp}(X)$ of $\mathcal{H}$ that is orthogonal to the kernel of X, $\ker(X)$.

For hermitian X with an eigendecomposition $X = \sum_{0 \le i < d} \lambda_i |\alpha_j\rangle\langle\alpha_j|$, $\mathrm{supp}(X) = \sum_{i:\lambda_i \neq 0} |\alpha_j\rangle\langle\alpha_j|$.

2.2.6 Transpose; Complex conjugate

Given a linear map $X \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$, its computational basis representation is $\sum_{j,k} \langle b|_\mathsf{B} X |a\rangle_\mathsf{A} |b\rangle_\mathsf{B} \langle a|_\mathsf{A}$. Its **transpose** (with respect to the computational basis) is the map $X^\mathrm{T} \in \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{A})$ defined by $\sum_{a,b} \langle b|_\mathsf{B} X |a\rangle_\mathsf{A} |a\rangle_\mathsf{A} \langle b|_\mathsf{B}$. The **complex conjugate** of X is the map $X^* \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$ defined by $\sum_{a,b} \langle b|_\mathsf{B} X |a\rangle_\mathsf{A}^* |b\rangle_\mathsf{B} \langle a|_\mathsf{A}$. Note that this definition is also basis dependent, in that it reflects the fact that we have declared the computational basis vectors are real. We have the following equations: (1) $(XY)^\mathrm{T} = Y^\mathrm{T} X^\mathrm{T}$; (2) $(XY)^* = X^* Y^*$ (3) $X^\dagger = (X^*)^\mathrm{T}$.

2.2.7 Trace

Definition 2. For any operator $X \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$, the trace of X is $\mathrm{Tr}X := \sum_{0 \le j < d_Q} \langle j|X|j\rangle$.

The trace is a linear function and it has the **cyclic property**:

$$\forall X \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B}), Y \in \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{A}) : \ \mathrm{Tr}XY = \mathrm{Tr}YX.$$

From the cyclic property it follows that $\mathrm{Tr}M^{-1}XM = \mathrm{Tr}X$ for any invertible M, so although we used the computational basis in the definition above, the trace is basis independent. Note that $\mathrm{Tr}X^\mathrm{T} = \mathrm{Tr}X$, and $\mathrm{Tr}X^\dagger = (\mathrm{Tr}X)^*$.

Proposition 3. $\forall\ |\psi\rangle, |\phi\rangle \in \mathcal{H}$, $\mathrm{Tr}|\psi\rangle\langle\phi| = \langle\phi|\psi\rangle$. (Proof is an easy exercise.)

2.2.8 The Hilbert-Schmidt inner product

We can use the trace to define an inner product for spaces of linear maps:

Definition 4 (Hilbert-Schmidt inner product)**.** For X, Y in $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$ let $\langle X, Y\rangle := \mathrm{Tr}X^\dagger Y$.

This is also an inner product on pairs of hermitian operators from the *real* vector space $\mathrm{Herm}(\mathcal{H}_\mathsf{Q})$.

2.2.9 Positivity of operators

We say an operator $X \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$ is **positive semi-definite** (or simply "positive") and write $X \ge 0$, if $\langle\psi|X|\psi\rangle \ge 0$ for all $|\psi\rangle \in \mathcal{H}_\mathsf{Q}$. Given $X, Y \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$ we write $X \ge Y$ iff $X - Y \ge 0$.

2.2.10 Isometries and unitaries

Definition 5 (Isometries)**.** A map V in $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$ which satisfies $\langle V|\psi\rangle, V|\phi\rangle\rangle = \langle|\psi\rangle, |\phi\rangle\rangle$ (equivalently $\langle\psi|V^\dagger V|\phi\rangle = \langle\psi|\phi\rangle$) for all $|\psi\rangle, |\phi\rangle \in \mathcal{H}_\mathsf{A}$, is called an *isometry.* V is an isometry iff $V^\dagger V = \mathbb{1}_\mathsf{A}$.

Remark 6. $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B})$ contains isometries iff $d_\mathsf{B} \ge d_\mathsf{A}$.

Definition 7 (Unitaries)**.** If $d_{\mathsf{B}} = d_{\mathsf{A}}$ then an isometry $U \in \mathcal{L}(\mathcal{H}_{\mathsf{A}}, \mathcal{H}_{\mathsf{B}})$ satisfies $U^\dagger U = UU^\dagger = \mathbb{1}_{\mathsf{A}}$, so $U^\dagger = U^{-1}$. Such an isometry is called *unitary.*

Proposition 8 (Unitary diagonalisation of unitary operators)**.** Any unitary operator $U \in \mathcal{L}(\mathcal{H}_{\mathsf{Q}})$ can be written

$$U = WDW^\dagger.$$

where $W = \sum_{0\le j<d_{\mathsf{Q}}} |w_j\rangle_{\mathsf{Q}}\langle j|_{\mathsf{Q}}$, $D = \sum_{0\le j<d_{\mathsf{Q}}} \lambda_j |j\rangle_{\mathsf{Q}}\langle j|_{\mathsf{Q}}$ where, for each j, $|w_j\rangle_{\mathsf{Q}}$ is a normalised eigenvector of U corresponding to eigenvalue λ_j. The set $\{|w_j\rangle_{\mathsf{Q}} : 0 \le j < d_{\mathsf{Q}}\}$ is an orthonormal basis for $\mathcal{H}_{\mathsf{Q}}$ and the eigenvalues all have modulus one $|\lambda_j| = 1$, so W and D are unitary.

2.3 Tensor products

Given spaces $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$, the tensor product space $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$ consists of finite linear combinations of the "product vectors" (or elementary tensors) $\{|\psi\rangle_{\mathsf{A}} \otimes |\phi\rangle_{\mathsf{B}} : |\psi\rangle_{\mathsf{A}} \in \mathcal{H}_{\mathsf{A}}, |\phi\rangle_{\mathsf{B}} \in \mathcal{H}_{\mathsf{B}}\}$ where the tensor product $\otimes$ of two vectors is bilinear:

1. $(|u\rangle_{\mathsf{A}} + |v\rangle_{\mathsf{A}}) \otimes |y\rangle_{\mathsf{B}} = |u\rangle_{\mathsf{A}} \otimes |y\rangle_{\mathsf{B}} + |v\rangle_{\mathsf{A}} \otimes |y\rangle_{\mathsf{B}}$,
2. $|u\rangle_{\mathsf{A}} \otimes (|x\rangle_{\mathsf{B}} + |y\rangle_{\mathsf{B}}) = |u\rangle_{\mathsf{A}} \otimes |x\rangle_{\mathsf{B}} + |w\rangle_{\mathsf{A}} \otimes |y\rangle_{\mathsf{B}}$,
3. $(\alpha|u\rangle_{\mathsf{A}}) \otimes |x\rangle_{\mathsf{B}} = \alpha(|u\rangle_{\mathsf{A}} \otimes |x\rangle_{\mathsf{B}})$,
4. $|u\rangle_{\mathsf{A}} \otimes (\alpha|x\rangle_{\mathsf{B}}) = \alpha(|u\rangle_{\mathsf{A}} \otimes |x\rangle_{\mathsf{B}})$.

for all $|u\rangle_{\mathsf{A}}, |v\rangle_{\mathsf{A}} \in \mathcal{H}_{\mathsf{A}}$, $|x\rangle_{\mathsf{B}}, |y\rangle_{\mathsf{B}} \in \mathcal{H}_{\mathsf{B}}$, $\alpha \in \mathbb{C}$. It has an inner product defined on product vectors by

$$\langle |u\rangle_{\mathsf{A}} \otimes |x\rangle_{\mathsf{B}}, |v\rangle_{\mathsf{A}} \otimes |y\rangle_{\mathsf{B}}\rangle = \langle |u\rangle_{\mathsf{A}}, |v\rangle_{\mathsf{A}}\rangle \langle |x\rangle_{\mathsf{B}}, |y\rangle_{\mathsf{B}}\rangle$$

and extended by linearity to the whole of $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$. If $\{|\alpha_j\rangle_{\mathsf{A}} : 0 \le j < d_{\mathsf{A}}\}$ is a basis for $\mathcal{H}_{\mathsf{A}}$, and $\{|\beta_k\rangle_{\mathsf{B}} : 0 \le k < d_{\mathsf{B}}\}$ is a basis for $\mathcal{H}_{\mathsf{B}}$, then the set $\{|\alpha_j\rangle_{\mathsf{A}} \otimes |\beta_k\rangle_{\mathsf{B}} : 0 \le j < d_{\mathsf{A}}, 0 \le k < d_{\mathsf{B}}\}$ is a basis for $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$, called the **product basis** of the two given bases. The **computational basis** for $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$ is the product basis of the computational bases for $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$, and it is orthonormal. The tensor product of $\mathcal{H}_{\mathsf{A}}$ with a one-dimensional space $\mathcal{H}_{\mathsf{W}}$ is isomorphic to $\mathcal{H}_{\mathsf{A}}$ and we identify $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{W}}$ with $\mathcal{H}_{\mathsf{A}}$ in the obvious way: $\forall z \in \mathbb{C}, |\psi\rangle \in \mathcal{H}_{\mathsf{A}}$, $|\psi\rangle_{\mathsf{A}} \otimes (z|0\rangle_{\mathsf{W}}) = z|\psi\rangle_{\mathsf{A}}$.

The tensor product of n vectors is multilinear in its n arguments, and $\mathcal{H}_{\mathsf{Q}_1} \otimes \cdots \otimes \mathcal{H}_{\mathsf{Q}_n}$ consists of finite linear combinations of product vectors $|\psi_1\rangle_{\mathsf{Q}_1} \otimes \cdots \otimes |\psi_n\rangle_{\mathsf{Q}_n}$. The tensor product is associative, so $(|a\rangle_{\mathsf{A}} \otimes |b\rangle_{\mathsf{B}}) \otimes |c\rangle_{\mathsf{A}} = |a\rangle_{\mathsf{A}} \otimes (|b\rangle_{\mathsf{B}} \otimes |c\rangle_{\mathsf{A}}) = |a\rangle_{\mathsf{A}} \otimes |b\rangle_{\mathsf{B}} \otimes |c\rangle_{\mathsf{A}} \in \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}} \otimes \mathcal{H}_{\mathsf{A}}$.

2.3.1 Tensor products of bras

The tensor product of the dual spaces for $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$ consists of finite linear combinations of product bras $\langle\psi| \otimes \langle\phi|$, where $\otimes$ is again bilinear. This space is isomorphic to the dual of $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$, and we make the identification $(|\psi\rangle_{\mathsf{A}} \otimes |\phi\rangle_{\mathsf{B}})^\dagger = \langle\psi|_{\mathsf{A}} \otimes \langle\phi|_{\mathsf{B}}$. This extends to the whole of $|\psi\rangle_{\mathsf{A}} \otimes |\phi\rangle_{\mathsf{B}}$ by conjugate linearity of the adjoint $\dagger$, e.g. if $|\chi\rangle_{\mathsf{AB}} = \sum_{j,k} c_{jk} |j\rangle_{\mathsf{A}} \otimes |k\rangle_{\mathsf{B}}$ then $\langle\chi|_{\mathsf{AB}} = \sum_{j,k} c^*_{jk} \langle j|_{\mathsf{A}} \otimes \langle k|_{\mathsf{B}}$

2.3.2 Tensor products of linear maps

Given $X \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A}), Y \in \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{B})$ we identify $X \otimes Y \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A}) \otimes \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{B})$ with the element of $\mathcal{L}(\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B})$ such that, for all $|\alpha\rangle_\mathsf{A}$, $|\beta\rangle_\mathsf{B}$

$$X \otimes Y|\alpha\rangle_\mathsf{A} \otimes |\beta\rangle_\mathsf{B} = (X|\alpha\rangle_\mathsf{A})_\mathsf{A} \otimes (Y|\beta\rangle_\mathsf{B})_\mathsf{B}.$$

This extends by linearity

$$\left(\sum_i \lambda_i X_i \otimes Y_i\right)\left(\sum_j \mu_j|\alpha_j\rangle_\mathsf{A} \otimes |\beta_j\rangle_\mathsf{B}\right) = \sum_i \sum_j \lambda_i\mu_j X_i \otimes Y_i|\alpha_j\rangle_\mathsf{A} \otimes |\beta_j\rangle_\mathsf{B}$$

to an isomorphism between $\mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A}) \otimes \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{B})$ and $\mathcal{L}(\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B})$.

Just like for vectors, given bases $\{X_i : i \in \{1, \ldots, d_\mathsf{A} \times d_\mathsf{A}\}\} \subset \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A}), \{Y_j : j \in \{1, \ldots, d_\mathsf{B} \times d_\mathsf{B}\}\} \subset \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{B})$, there is a product basis $\{X_i \otimes Y_j\} \subset \mathcal{L}(\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}) \otimes \mathcal{L}(\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B})$. In terms of computational bases, the isomorphism mentioned amounts to the identification

$$|c\rangle_\mathsf{A}\langle a|_\mathsf{A} \otimes |d\rangle_\mathsf{B}\langle b|_\mathsf{B} = |c\rangle_\mathsf{A} \otimes |d\rangle_\mathsf{B}\langle a|_\mathsf{A} \otimes \langle b|_\mathsf{B}.$$

Since a bra $\langle\psi|_\mathsf{Q}$ is a linear map from $\mathcal{H}_\mathsf{Q}$ to the one-dimensional space $\mathbb{C}$, $\langle\psi|_\mathsf{Q}\otimes X$ belongs to $\mathcal{L}(\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{Q}, \mathcal{H}_\mathsf{A})$. In a tensor product between X and a vector $|\psi\rangle_\mathsf{Q}$ in $\mathcal{H}_\mathsf{Q}$, we regard the vector as the linear map $|\psi\rangle_\mathsf{Q} : z \mapsto |\psi\rangle_\mathsf{Q}$ in $\mathcal{L}(\mathbb{C}, \mathcal{H}_\mathsf{Q})$. So $|\psi\rangle_\mathsf{Q} \otimes X \in \mathcal{L}(\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{Q})$.

2.3.3 Matrix representation of tensor products

We choose the matrix representation of the computational basis of a composite system so that the tensor product of two or more objects corresponds to the “Kronecker product” of the corresponding matrices: For example, if

$$\begin{aligned}
|a\rangle_\mathsf{A} =&\alpha_0|0\rangle_\mathsf{A} + \alpha_1|1\rangle_\mathsf{A} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix},\ |b\rangle_\mathsf{B} = \beta_0|0\rangle_\mathsf{B} + \beta_1|1\rangle_\mathsf{B} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},\\
|c\rangle_\mathsf{A} =&\gamma_0|0\rangle_\mathsf{A} + \gamma_1|1\rangle_\mathsf{A} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix},\ L \in \mathcal{L}(\mathcal{H}_\mathsf{B}, \mathcal{H}_\mathsf{A}) = \sum_{0\le j,k<2} \lambda_{jk}|j\rangle_\mathsf{A}\langle k|_\mathsf{B} = \begin{pmatrix} \lambda_{00} & \lambda_{01} \\ \lambda_{10} & \lambda_{11} \end{pmatrix},\\
M \in&\mathcal{L}(\mathcal{H}_\mathsf{A}) = \sum_{0\le j,k<2} \mu_{jk}|j\rangle\langle k|_\mathsf{A} = \begin{pmatrix} \mu_{00} & \mu_{01} \\ \mu_{10} & \mu_{11} \end{pmatrix}
\end{aligned}$$

then

$$|a\rangle_\mathsf{A} \otimes |b\rangle_\mathsf{B} = \begin{pmatrix} \alpha_0\beta_0 \\ \alpha_0\beta_1 \\ \alpha_1\beta_0 \\ \alpha_1\beta_1 \end{pmatrix},\ |a\rangle_\mathsf{A} \otimes |b\rangle_\mathsf{B} \otimes |c\rangle_\mathsf{A} = \begin{pmatrix} \alpha_0\beta_0\gamma_0 \\ \alpha_0\beta_0\gamma_1 \\ \alpha_0\beta_1\gamma_0 \\ \alpha_0\beta_1\gamma_1 \\ \alpha_1\beta_0\gamma_0 \\ \alpha_1\beta_0\gamma_1 \\ \alpha_1\beta_1\gamma_0 \\ \alpha_1\beta_1\gamma_1 \end{pmatrix},\ |a\rangle \otimes L = \begin{pmatrix} \lambda_{00}\alpha_0 & \lambda_{01}\alpha_0 \\ \lambda_{10}\alpha_0 & \lambda_{11}\alpha_0 \\ \lambda_{00}\alpha_1 & \lambda_{01}\alpha_1 \\ \lambda_{10}\alpha_1 & \lambda_{11}\alpha_1 \end{pmatrix},$$

$$L\otimes M = \begin{pmatrix} \lambda_{00}\mu_{00} & \lambda_{00}\mu_{01} & \lambda_{01}\mu_{00} & \lambda_{01}\mu_{01} \\ \lambda_{00}\mu_{10} & \lambda_{00}\mu_{11} & \lambda_{01}\mu_{10} & \lambda_{01}\mu_{11} \\ \lambda_{10}\mu_{00} & \lambda_{10}\mu_{01} & \lambda_{11}\mu_{00} & \lambda_{11}\mu_{01} \\ \lambda_{10}\mu_{10} & \lambda_{10}\mu_{11} & \lambda_{11}\mu_{10} & \lambda_{11}\mu_{11} \end{pmatrix},\ \langle a|\otimes L = \begin{pmatrix} \alpha_0^*\lambda_{00} & \alpha_0^*\lambda_{01} & \alpha_1^*\lambda_{00} & \alpha_1^*\lambda_{01} \\ \alpha_0^*\lambda_{10} & \alpha_0^*\lambda_{11} & \alpha_1^*\lambda_{10} & \alpha_1^*\lambda_{11} \end{pmatrix}.$$

2.4 Probability notation

1. If X is a random variable (RV) which takes values in $\mathcal{A}_X$, then we will use P_X, Q_X etc. for probability distributions on $\mathcal{A}_X$, which for our (finite/discrete) purposes are really *probability mass functions* $P_X : \mathcal{A}_X \to [0,1]$ which satisfy $\sum_{x \in \mathcal{A}_X} P_X(x) = 1$. I will call them 'distributions'.

2. To say "X is distributed according to P_X" is to say that $\Pr(X = x) = P_X(x)$. But note that this equation is not the *definition* of P_X: For example, $P_X(X) \neq \Pr(X = X)$. $P_X(X)$ is the function P_X applied to the RV X. Therefore, $P_X(X)$ is itself an RV which takes the value $P_X(x)$ with probability $P_X(x)$. $\Pr(X = X)$ is just the number one! If X is distributed according to P_X, then $Q_X(X)$ is a RV which takes the value $Q_X(x)$ with probability $P_X(x)$.

3. A joint distribution P_{XY} for two random variables X, Y is a function $P_{XY} : \mathcal{A}_X \times \mathcal{A}_Y \to [0,1]$ satisfying $\sum_{x,y} P_{XY}(x, y) = 1$. If we say "the (joint) distribution of X, Y is P_{XY}" this means $\Pr(X = x, Y = y) = P_{XY}(x, y)$. A conditional distribution $P_{X|Y}$ is a function $P_{X|Y} : \mathcal{A}_X \times \mathcal{A}_Y \to [0,1]$ satisfying $\sum_x P_{X|Y}(x|y) = 1$ for all $y \in \mathcal{A}_Y$. $P_{Y|X=x} : \mathcal{A}_Y \to [0,1]$ denotes a distribution for Y conditioned on $X = x$.

4. Distributions and conditional distributions with the same label but different subscripts are assumed to be compatible according to the rules of probability. So, if we have a joint distribution Q_{XY} then Q_Y must be the marginal distribution of Y when the joint distribution of X, Y is Q_{XY} (i.e. $Q_Y(y) = \sum_x Q_{XY}(x, y)$) and similarly for Q_X. Likewise, the product rule must hold $Q_{XY}(x, y) = Q_{X|Y}(x|y)Q_Y(y) = Q_{Y|X}(y|x)Q_X(x) = Q_{Y|X=x}(y)Q_X(x)$.

5. The 'default' distribution of everything is called P: That is, unless otherwise stated, $\Pr(X = x, Y = y, \dots, W = w) = P_{XY\cdots W}(x, y, \dots, w)$.

3 Postulates of Quantum Mechanics

This section reviews the postulates of quantum mechanics as given in many introductory quantum mechanics courses, but in a slightly unusual form. In the course we will build on these to obtain more general notions of state, measurement and time evolution.

3.1 States

Postulate 9 (State postulate)**.** To any system Q there is associated a complex Hilbert space $\mathcal{H}_{\mathsf{Q}}$ (the 'state space' of Q) and the state of the system is described by a **state vector** $|\psi\rangle$ which is simply a unit vector in $\mathcal{H}_{\mathsf{Q}}$ i.e. $\langle\psi|\psi\rangle = 1$. State vectors $|\psi\rangle$ and $|\psi'\rangle$ which differ only by an global phase, i.e. $|\psi'\rangle = e^{i\alpha}|\psi\rangle$ for some $\alpha \in \mathbb{R}$, represent the same state of the system.

It is important to note that *relative* phase in a superposition (a linear combination) does have physical significance, e.g. the state vectors $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle + e^{i\pi}|1\rangle)/\sqrt{2}$ are certainly not equivalent up to a global phase and so do not represent the same physical state (in fact, they are orthogonal).

The smallest interesting systems are two-dimensional, and we call these **qubits**. An arbitrary state vector of a qubit can be written thus

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}.$$

where the normalisation condition is $\langle\psi|\psi\rangle = |\alpha_0|^2 + |\alpha_1|^2 = 1$. Since,

$$|\psi\rangle = e^{i\arg(\alpha_0)}\big(|\alpha_0||0\rangle + |\alpha_1|e^{i\phi}|1\rangle\big) = e^{i\arg(\alpha_0)}\big(\cos(\theta/2)|0\rangle + \sin(\theta/2)e^{i\phi}|1\rangle\big)$$

where $\phi = (\arg(\alpha_1) - \arg(\alpha_0))$ and $\theta/2 = \arccos|\alpha_0|$, we can parameterise the physically distinct state vectors by $\theta \in [0,\pi]$ and $\phi \in [0, 2\pi)$. Interpreting these as spherical polar coordinates lets us identify physically distinct state vectors with points on the surface of a three-dimensional sphere which, in this context, is called the **Bloch sphere**. Examples of physical qubits are photon polarization and electron spin.

3.2 Measurements

Postulate 10 (Measurement postulate)**.** A measurement on a system Q whose result X (a random variable) takes values in $\mathcal{A}_X$ is represented by a PVM E which assigns to each $x \in \mathcal{A}_X$ a projector $E(x)$ on $\mathcal{H}_{\mathsf{Q}}$ such that

1. Projectors for different values project onto orthogonal subspaces:

$$E(x)E(y) = \begin{cases} E(x) \text{ if } x = y, \\ 0 \text{ if } x \neq y. \end{cases}$$

2. The projectors sum to the identity operator $\sum_{x\in\mathcal{A}_X} E(x) = \mathbb{1}_{\mathsf{Q}}$.

If the state vector of Q is $|\psi\rangle$, then for a measurement on Q with PVM E:

1. The probability that the measurement result is x is
$$\Pr(X = x) = \langle\psi|E(x)|\psi\rangle.$$

2. Immediately after the measurement, conditioned $X = x$, the state vector of Q becomes
$$\frac{E(x)|\psi\rangle}{\|E(x)|\psi\rangle\|}.$$

In the case where $\mathcal{A}_X$ is a set of real numbers, the eigendecomposition of hermitian operators places the PVMs on Q in one-to-one correspondence with hermitian operators on $\mathcal{H}_{\mathsf{Q}}$, which in this context are called *observables.*

3.3 Time evolution

The dynamics of a closed quantum system Q are encoded by a *Hamiltonian* H which is a an hermitian operator on $\mathcal{H}_{\mathsf{Q}}$. Its physical meaning as an observable is that it measures the total energy of the system.

Postulate 11 (Time evolution)**.** The time evolution of the state vector of a closed quantum system is given by a linear differential equation called the **Schrödinger equation**:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle.$$

Here, $\hbar$ is a physical constant called the *reduced Planck constant.* We will choose our units so that it's equal to one. Solving the Schrödinger equation we obtain, for any $|\psi(t_1)\rangle$,

$$|\psi(t_2)\rangle = U(\Delta t)|\psi(t_1)\rangle$$

where $U(\Delta t) := \exp(-iH\Delta t)$ is a *unitary* operator on $\mathcal{H}_{\mathsf{Q}}$, and $\Delta t := t_2 - t_1$.

We won't talk about Hamiltonians again in this course. For us the important point is that the time evolution of a closed system is unitary and, furthermore, for *any* unitary transformation U we can find a Hamiltonian H and time interval Δt such that

$$U = \exp(-iH\Delta t).$$

3.4 Composite systems

Postulate 12 (Composite systems)**.** If systems Q_j, $1 \le j \le n$, have state spaces $\mathcal{H}_{\mathsf{Q}_j}$ then the state space of the composite system $\mathsf{Q}_1\mathsf{Q}_2 \cdots \mathsf{Q}_n$ is the tensor product $\mathcal{H}_{\mathsf{Q}_1} \otimes \mathcal{H}_{\mathsf{Q}_2} \otimes \cdots \otimes \mathcal{H}_{\mathsf{Q}_n}$. If the state vector of Q_j is $|\psi_j\rangle_{\mathsf{Q}_j}$ for each j then the state vector of the composite system is the product vector $|\psi_1\rangle_{\mathsf{Q}_1} \otimes |\psi_2\rangle_{\mathsf{Q}_2} \otimes \cdots \otimes |\psi_n\rangle_{\mathsf{Q}_n}$.

Any state vector of the composite system which is *not* a product vector is called *entangled.*

Measurements on composite systems

Proposition 13. Given a measurement on system A represented by a PVM E_{A}, the corresponding PVM on AB is the one which associates outcome x to $E(x)_{\mathsf{A}} \otimes \mathbb{1}_{\mathsf{B}}$.

♣♣Use the composite systems postulate and the measurement postulate to derive this proposition.