Quantum Information & Computation

Adam Kelly

February 4, 2021

This set of notes is a work-in-progress account of the course 'Quantum Information & Computation', originally lectured by Prof Richard Jozsa in Lent 2020 at Cambridge. These notes are not a direct transcription of the lectures, but they do roughly follow what was lectured (in content and in structure).

These notes are likely to be more succinct than other lecture notes of mine, and I have left out various aspects of what was taught. If you spot any errors in this set of notes, I can be contacted at ak2316@cam.ac.uk.

Contents

1 Principles of Quantum Mechanics

1.1 Dirac Notation

Let V be a finite dimensional complex vector space with a (hermitian) inner product. In Dirac notation, we write vectors as $|v\rangle$ called *ket vectors*.

We will often work with two dimensional space V_2 with a chosen orthonormal basis $\{|0\rangle, |1\rangle\}$, labelled by bit values.

By convention, kets are always written as column vectors in components. For example,

$$|v\rangle = a\,|0\rangle + b\,|1\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \qquad a, b \in \mathbb{C}.$$

The *conjugate transpose* of $|v\rangle$ is a *bra vector*, written in mirror image notation,

$$\langle v| = |v\rangle^\dagger = a^*\,\langle 0| + b^*\,\langle 1| = \begin{pmatrix} a^* & b^* \end{pmatrix},$$

and bras are always written as row vectors in components.

More formally, bra vectors $\langle v|$ is an element of the duel vector space V^* of V under the canonical isomorphism $V \cong V^*$, given by the inner product. That is, $\langle v|$ is a linear map $|w\rangle \mapsto$ the inner product of $|v\rangle$ with $|w\rangle$. Note that this inner product is linear in $|w\rangle$ and antilinear in $|v\rangle$ (linear in $\langle v|$).

If $|w\rangle = c\,|0\rangle + d\,|1\rangle$, then the inner product of $|v\rangle$ with $|w\rangle$ is written by juxtaposing the bra and ket,

$$\langle v|w\rangle = |v\rangle^\dagger\,|w\rangle = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = a^*c + b^*d.$$

For example, the orthonormality of basis vectors can be written as $\langle i|j\rangle = \delta_{ij}$.

1.2 Tensor Products of Vectors

For some vector space V of dimension m on a basis $|e_1\rangle, \ldots, |e_m\rangle$, and another vector space V of dimension n on a basis $|f_1\rangle, \ldots, |f_n\rangle$, the *tensor product space* $V \otimes W$ has dimension mn with orthonormal basis $\{|e_i\rangle \otimes |f_j\rangle\}$, $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, where $\otimes$ is bilinear. Then a general ket vector in $V \otimes W$ is

$$|v\rangle = \sum c_{ij}\,|e_i\rangle \otimes |f_j\rangle\,.$$

There is a natural *bilinear* map $f : V \times W \to V \otimes W$. If $|\alpha\rangle = \sum a_i\,|e_i\rangle$ and $|\beta\rangle = \sum b_j\,|f_j\rangle$, then

$$\begin{aligned}(|\alpha\rangle, |\beta\rangle) \mapsto f \mapsto |a\rangle \otimes |b\rangle \\ &= \left(\sum a_i\,|e_i\rangle\right) \otimes \left(\sum b_j\,|f_j\rangle\right) \\ &= \sum_{ij} a_i b_j\,|e_i\rangle \otimes |f_j\rangle\,.\end{aligned}$$

Note that $\otimes$ is not commutative. For example, if $V = W$ then $|\alpha\rangle\,otimes\,|\beta\rangle \neq |\beta\rangle \otimes |\alpha\rangle$ in general. We will often omit the symbol $\otimes$ and will write $|\alpha\rangle \otimes |\beta\rangle$ as $|\alpha\rangle\,|\beta\rangle$.

The mapping f is *not surjective*. With this in mind, we introduce the notions of *product* and *entangled* vectors.

Definition 1.1 (Product and Entangled)**.** Any $|\xi\rangle \in V \otimes W$ of the form $|\xi\rangle = |\alpha\rangle \otimes |\beta\rangle$ is called a *product vector*. Any $|\xi\rangle$ that is *not* a product vector is called *entangled*.

We will mostly be concerned with tensor products of the 2 dimensional V_2 with itself (possible many times over). For the k-fold tensor power, we write $\bigotimes^k V_2 = V_2 \otimes \cdots \otimes V_2$. This has dimension 2^k and orthonormal basis

$$|i_1\rangle \otimes \cdots \otimes |i_k\rangle, \qquad i_1, \ldots, i_k \in \{0, 1\}.$$

These basis vectors are labelled by 2^k k-bitstrings. We will often write $|i_1\rangle \otimes \cdots |i_k\rangle$ as $|i_1\rangle \cdots |i_k\rangle$ or $|i_1 \ldots i_k\rangle$.

Example 1.2. The vector $|v\rangle = |00\rangle + |11\rangle$ in $V_2 \otimes V_2$ is *entangled*. To see this, suppose we could write $|v\rangle$ as a product:

$$\begin{aligned} |v\rangle &= (a\,|0\rangle + b\,|1\rangle) \otimes (c\,|0\rangle + d\,|1\rangle) \\ &= ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle. \end{aligned}$$

Comparing with the coefficients of $|v\rangle$, we get $ad = 1$, $ad = 0$, $bc = 0$ and $ad = 1$. But then $abcd = (ac)(bd) = 1$ and $abcd = (ad)(bc) = 0$, which is a contradiction. Thus $|v\rangle$ must be entangled.

We can show that

$$|v\rangle = a_{00}\,|00\rangle + a_{01}\,|01\rangle + a_{10}\,|10\rangle + a_{11}\,|11\rangle$$

is entangled if and only if $\det(a_{ij}) \neq 0$. For general dimensions,

$$\sum_{i=1,j=1}^{m,n} A_{ij}\,|i\rangle\,|j\rangle$$

is a product vector if and only if the matrix $[A_{ij}]$ has rank 1.

The inner product on $V \otimes W$ is induced the inner products on V and W, 'applied slotwise'. For product states $|\alpha_1\rangle\,|\beta_2\rangle$ and $|\alpha_2\rangle\,|\beta_2\rangle$, the inner product is

$$(\langle b_1|\,\langle\alpha_1|)(|\alpha_2\rangle\,|\beta_2\rangle) = \langle\alpha_1|\alpha_2\rangle\,\langle\beta_1|\beta_2\rangle,$$

and we extend by linearity to all $|\xi\rangle \in V \otimes W$.

Remark (Notation)**.** For the bra vector of $|\alpha\rangle\,|\beta\rangle$, we often reverse the order and write is as $\langle\beta|\,\langle\alpha|$. It is always important to keep track of component slots. We sometimes include explicit labels (for example, $|\alpha\rangle_A\,|\beta\rangle_B$ has bra vector ${}_A\langle\alpha|_B\,|\beta\rangle = {}_B\langle\beta|\,{}_A\langle\alpha|$).

1.3 Quantum Principles

We will now state some axioms that describe quantum mechanics.

1.3.1 QM1 – Physical States

Axiom (QM1)**.** The state of any (isolated) physical system S are represented by unit vectors in a complex vector space V with a given inner product.

The simplest nontrivial case is $V = V_2$, the two dimensional complex vector space. We choose a pair of orthonormal vectors $|0\rangle$ and $|1\rangle$. Then a general state is

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle, \qquad \alpha, \beta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 = 1.$$

We say $|\psi\rangle$ is a *superposition* of $|0\rangle$ and $|1\rangle$ with *amplitudes* α and β respectively.

Definition 1.3 (Qubit)**.** A *qubit* is any quantum system with a two dimensional state space and a chosen orthonormal basis labelled $|0\rangle$, $|1\rangle$ called the *computational* basis, *standard* basis or Z-basis.

Definition 1.4 (Conjugate Basis for a Qubit)**.** Given an orthonormal pair $|0\rangle$ and $|1\rangle$, we get the orthonormal pair

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), |-\rangle \qquad = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right).$$

We call this the *conjugate* basis or X-basis.

1.3.2 QM2 – Composite Systems

Axiom (QM2)**.** If system S_1 has state space V_1 and system S_2 has state space V_2 then the joint system S_1S_2, obtained by taking S_1 and S_2 together, has state space $V_1 \otimes V_2$.

Comparing this axiom with the classical analog, the corresponding statement has a Cartesian product rather than a tensor product. So giving a state for a classical S_1S_2 is just giving a state for S_1 and giving a state for S_2. Thus the dimension of the system grows linearly with the number of systems, whereas in the quantum sense (with this axiom), the composite system grows *exponentially* with the number of systems.

Example 1.5 (n-qubit system)**.** An n-qubit system has as state space $V_2^{\otimes n}$, with dimension 2^n. It has the computational basis $|i_1 \dots i_n\rangle$, labelled by all 2^n n-bit string.

An n qubit state $|\psi\rangle$ is a *product state* if is the tensor product of n single qubit states $|\psi\rangle = |v_1\rangle |v_2\rangle \cdots |v_n\rangle$, otherwise it is *entangled*.

Before we state the next axiom, we need to introduce how to work with linear maps in Dirac notation.

Consider linear maps on V_2 (higher-dimensions are similar) with $|v\rangle = a|0\rangle + b|1\rangle$ and $|w\rangle = c|0\rangle + d|1\rangle$. Then the 'ket-bra' product is

$$M = |v\rangle\langle w| = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} c^* & d^* \end{pmatrix} = \begin{pmatrix} ac^* & ad^* \\ bc^* & bd^* \end{pmatrix},$$

which is a linear map on V_2. For any $|x\rangle = (x_0, x_1)^T$, we have

$$M|x\rangle = \left(|v\rangle\langle w|\right)|x\rangle = |v\rangle \langle w|x\rangle .$$

These are all rank 1 mappings, and thus is not the most general linear map.

For a general linear map say

$$A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix},$$

we note that for basis states,

$$|0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad |1\rangle\langle 0| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

and we can see that these are a basis for 2×2 matrices, and we can express

$$A = a_{00} |0\rangle\langle 0| + a_{01} |0\rangle\langle 1| + a_{10} |1\rangle\langle 0| + a_{11} |1\rangle\langle 1| = \sum a_{ij} |i\rangle\langle j| .$$

Another useful property comes from the cycic property of the trace:

$$\langle v|w\rangle = \mathrm{tr}(|v\rangle\langle w|).$$

An important operation comes from the above when $|v\rangle = |w\rangle$ (in any dimension), and where $|v\rangle$ is normalized so $\langle v|v\rangle = 1$. Then $\Pi_v = |v\rangle\langle u|$ is the operation of projection onto the 1-dimensional subspace spanned by $|v\rangle$. For example, we have

$$\Pi_v \Pi_v = |v\rangle \langle v|v\rangle = |v\rangle\langle v| = \Pi_v.$$

More generally if E is any d-dimensional linear subspace of n-dimensional space V and $\{|e_1\rangle, \ldots, |e_d\rangle\}$ is any orthonormal basis of E, then $\Pi_E = |e_1\rangle\langle e_1| + \cdots + |e_d\rangle\langle e_d|$ is the operation of projection on to E. Π_E is independent of the orthonormal basis chosen.

We also want to have tensor products of maps. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} p & q \\ v & s \end{pmatrix}$$

are linear maps on V_2 then $A \otimes B : V_2 \otimes V_2 \to V_2 \otimes V_2$ defined by the basis action

$$|i\rangle |j\rangle \longmapsto (A|i\rangle)(B|j\rangle),$$

with linear extension.

For example, for any product sate $|v\rangle |w\rangle$, we have $(A \otimes B) |v\rangle |w\rangle = (A|v\rangle)(B|w\rangle)$.

In components for the 2×2 case we have

$$A \otimes B = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix} = \left(\begin{array}{cc|c} ap & aq & bB \\ av & as & \\ \hline & cB & dB \end{array}\right),$$

with other cases done similarily.

Some important special cases are $I \otimes A$ and $A \otimes I$, the action of A on the second (or first) component space of $V_2 \otimes V_2$. We often refer to these as *local operations* on subsystems of composite systems.

1.3.3 QM3 – Physical Evolution of Quantum Systems

Axiom (QM3)**.** Any physical (finite time) evolution of a quantum system is represented by a *unitary* linear operation on the vector space of states.

The analog to this from the Quantum Mechanics course would be Schrödinger's equation and specifically is the finite dimensional version of Schrödinger's equation, where instead of a hamiltonian we have a hermitian operation on the state space, where evolution is

$$\frac{\mathrm{d}|\psi\rangle}{\mathrm{d}t} = H|\psi\rangle$$

If H is time independent, then $|\psi\rangle = e^{-\frac{i}{\hbar}Ht}|\psi_0\rangle$, which is the matrix exponential. This also implies that A hermitian gives e^{iA} unitary.

Recall that a matrix U is unitary if $U^{-1} = U^\dagger$, that is, if and only if U maps any orthonormal basis to an orthonormal set of vectors, which occurs if and only if the columns or rows of the matrix of U form an orthonormal set of vectors.

The last piece of notation we will introduce is the partial inner products. For vectors in $V \otimes W$, then any ket $|v\rangle \in V$ defines a linear map $V \otimes W \to W$ called the 'partial inner product with $|v\rangle$', defined on basis vectors $|e_i\rangle|f_j\rangle \in V \otimes W$ where $|e_i\rangle|f_j\rangle \mapsto \langle v|e_i\rangle|f_j\rangle$, and similarity for $|w\rangle \in W$ giving a map $V \otimes W \to V$). If $V = W$ (as often occurs), it is important to specify which space of $V \otimes V$ is being used.

Example 1.6. For $V = V_2$, and $|\xi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle \in V \otimes V$, we can form the partial inner product with $|0\rangle$ on *either* space. So first label the spaces as $V_A \otimes V_B$ where $V_A = V_B = V$. Then the orthonormality relations $\langle i|j\rangle = \delta_{ij}$ gives

$$\begin{aligned} {}_A\langle 0|\xi\rangle_{AB} &= a_A \langle 0|0\rangle_A |0\rangle_B + b_A \langle 0|0\rangle_A |1\rangle_B + c_A \langle 0|1\rangle_A |0\rangle_B + d_A \langle 0|1\rangle_A |1\rangle_B \\ &= a|0\rangle_B + b|1\rangle_B . \end{aligned}$$

That is, we pick out the terms of $|\xi\rangle_{AB}$ with a zero in the A slot.

1.3.4 QM4 – Quantum Measurement & Born Rule

This axiom describes how classical information is extracted from quantum states. Fundamentally there is a physical limitation on this process and it is rather different than the extraction of classical information.

Typically in a classical setting, we assume that measurements can occur without changing the physical system being studied. For example, if you measured the height of a tree, you wouldn't expect the act of measuring to make the tree any smaller. This is different in a quantum setting, as we shall see.

We will work with a given *single* instance of a quantum state $|\psi\rangle$ of a physical system, with a state space V of dimension n.

We first have the *basic Born rule*, a complete projective measurement (or von Neumann measurement). Let $B = \{|e_1\rangle, \ldots, |e_n\rangle\}$ be any orthonormal basis of V. We can make a measurement on $|\psi\rangle$ relative to the basis B.

Letting $|\psi\rangle = \sum a_i |e_i\rangle$, the possible measurement outcomes are $j = 1, 2, \ldots, n$, and the probability of getting an outcome j is $\text{prob}(j) = |\langle e_j|\psi\rangle|^2 = |a_j|^2$. If outcome j is seen, then the state is no longer $|\psi\rangle$, but it has been *collapsed* to $|\psi_j\rangle = |e_j\rangle$.

Clearly this transformation is not unitary, and thus we have a completely different physical process to evolution as described before. Notably, repeated measurement gives only the same result, with certainty. We would not get further samples of the $|a_j|^2$ distribution.

To rephrase, the probability is the squared length of $|\psi\rangle$ onto the basis direction $|e_j\rangle$, and the post measurement state is that projection, re-normalized to unit length.

The 'complete' above refers to the *one*-dimensionality of the orthonormal subspaces, defined by basis states. We can similarity describe incomplete projective measurements.

Let $\{E_1, E_2, \ldots, E_d\}$ by a decomposition of V into d mutually orthogonal subspaces, so $V = E_1 \oplus \cdots \oplus E_d$, and let Π_i be the projection onto E_i. Then the *incomplete* projective measurement of $|\psi\rangle$ with respect to the orthogonal decomposition is the following quantum operation. The possible outcomes are $j = 1, \ldots, d$, and prob(j) = squared length of projection of $|\psi\rangle$ into $E_j = \langle\psi|\Pi_j^\dagger \Pi_j|\psi\rangle = \langle\psi|\Pi_j|\psi\rangle$. (Recall that $\Pi\Pi = \Pi$ and $\Pi^\dagger = \Pi$). This is also $\mathrm{tr}(\Pi_j |\psi\rangle\langle\psi|)$.

The post measurement state is the projected 'collapsed' vector, re-normalised. So $|\psi_j\rangle = \frac{\Pi_j|\psi\rangle}{\mathrm{prob}(j)}$.

Example 1.7 (Parity Measurement on 2-Qubits)**.** The parity of a 2-bit string $b_1 b_2$ is the modulo 2 sum $b_1 \oplus b_2$. The parity measurement on two qubits is teh incomplete measurement with orthogonal decomposition

$$E_0 = \mathrm{span}\{|00\rangle, |11\rangle\}, \qquad E_1 = \mathrm{span}\{|01\rangle, |10\rangle\}.$$

For $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, we see 0 with probability $|a|^2 + |d|^2$, and the post measurement state would be $\frac{a|00\rangle + d|11\rangle}{\sqrt{|a|^2+|d|^2}}$.

In some texts (especially pre-2000s) they will typically introduce measurement in the context of 'quantum observables'.

Definition 1.8 (Observable)**.** A quantum observable θ is a hermitian operator on V so that θ has real eigenvalues λ_j for $j = 1, \ldots, d$, and orthogonal eigenspaces Λ_j, with $\dim(\Lambda_j)$ being the multiplicity of λ_j, and $V = \Lambda_1 \oplus \cdots \oplus \Lambda_d$.

The measurement of a quantum observable θ is then just the incomplete measurement relative to this orthogonal decomposition. However, these observables' eigenvalues relate to physical properties, and we thus normally label the outcomes by their eigenvalues λ_j, not just j as before. We also have things like $\langle\theta\rangle = \sum \lambda_j \,\mathrm{prob}(j)$, which obviously depend on the eigenvalues.

But for the purposes of providing information about $|\psi\rangle$ and its post measurement state, the choice of labelling is immaterial and thus we won't be that interested in observables.

A special case of incomplete measurement relates to measuring a component part of a quantum system, the *extended Born rule*.

Suppose $|\psi\rangle$ is a state of a composite system S_1S_2 with state space $V \otimes W$. Let $B_V = \{|e_1\rangle, \ldots, |e_n\rangle\}$ be an orthonormal basis of V. We can uniquely express $|\psi\rangle$ uniquely as $|\psi\rangle_{VW} = \sum_i |e_i\rangle_V |\xi_i\rangle_W$, where $|\psi_i\rangle$ are generally not normalized, and not orthogonal. In fact, $|\xi_i\rangle_W = {}_V\langle e_i|\psi\rangle_{VW}$. Thus $|\psi\rangle$ normalised implies $\sum_{i=1}^n \langle\xi_i|\xi_i\rangle = 1 = \langle\psi|\psi\rangle$.

Now, if we say perform a measurement of $|\psi\rangle$ relative to a basis B_V of V (a 'complete measurement on V but not $V \otimes W$') with outcomes $i = 1, \ldots, n$, and corresponding orthogonal subspaces (of $V \otimes W$), $E_i = \mathrm{span}\{|e_i\rangle \otimes |\xi\rangle \mid |\xi\rangle \in W\}$,

which we write as $|e_i\rangle \otimes W$. The corresponding projectors are $\Pi_i = |e_i\rangle\langle e_i| \otimes I_W$, and $\text{prob}(i) = \langle \xi_i | \xi_i \rangle$. The post measurement state for i is $|\psi_i\rangle = \frac{|e_i\rangle_V |\xi_i\rangle_w}{\sqrt{\langle \xi_i | \xi_i \rangle}}$.

Remark. According to QM4 (all of the measurement rules), two different states with guaranteed (probability 1) different outcomes for some measurement, must lie in orthogonal subspaces (they must be orthogonal themselves). So non-orthogonal state, although physically different, can never be reliably distinguished ('as information') by any quantum process.

Also, if $|\psi\rangle$ has dimension n, any measurement on it has at most n outcomes. However, we can get more outcomes by adjoining an ancilla $|A\rangle$ of dimension m (independent of $|\psi\rangle$) and measure $|\psi\rangle |A\rangle$ to get up to mn outcomes.

1.4 Quantum Gates

We will finish looking at the principles of quantum mechanics that we will be using by looking at some basic unitary operations for qubits, which are known as *quantum gates*.

1.4.1 Single Qubit Gates

We will begin by listing some gates that act on a single qubit.

- The Hadamard gate,
$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$
We have $H|0\rangle = |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $H|1\rangle = |-\rangle$.

- The Pauli operations/gates,
$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
These matrices are all unitary and hermitian. These matrices have multiplicative (group) properties, for example $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$, so the eigenvalues are ± 1, and they all anti-commute, $\sigma_x \sigma_y = -\sigma_y \sigma_z$, and there's cyclic shifts, for example $\sigma_y \sigma_z = \sigma_x$.

We will often use the 'real versions' as quantum gates,
$$X = \sigma_x, \quad Z = \sigma_z, \quad Y = i\sigma_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$
Note that $X|0\rangle = |1\rangle$, so this gate is often known as the 'quantum NOT gate'. The X eigenbasis is $|+\rangle, |-\rangle$. The Z gate satisfies $Z|0\rangle = |0\rangle$, and $Z|1\rangle = -|1\rangle$. The Z eigenbasis is $|0\rangle$, and $|1\rangle$. Note the relation to the X and Z basis mentioned before.

- The phase gates
$$P(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}.$$
Note that $Z = P(\pi)$.

1.4.2 Two Qubit Gates

There are also gates that act on multiple qubits. The most common examples of two qubit gates are the 'controlled' gates.

- The CX, or controlled-X gate. This is a two qubit gate where the action of the X gate on the second qubit is controlled by the value of the first qubit.

  To define it on a basis, $CX\left|j\right\rangle\left|k\right\rangle = \left|j\right\rangle\left|j \oplus k\right\rangle = \left|j\right\rangle X^j\left|k\right\rangle$, where $\oplus$ is addition modulo 2. So $CX\left|0\right\rangle\left|\alpha\right\rangle = \left|0\right\rangle\left|\alpha\right\rangle = CX\left|1\right\rangle\left|\alpha\right\rangle = \left|1\right\rangle X\left|\alpha\right\rangle$.

  The matrix for the CX gate is

  $$CX = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

  We typically call the first qubit the *control* qubit and the second qubit the *target* qubit. The two qubits have asymmetrical roles, so sometimes we write CX_{ab} to mean a is the control qubit and b is the target qubit to make this explicit.

- The CZ, or controlled-Z gate. This is $CZ_{12}\left|i\right\rangle\left|j\right\rangle = \left|j\right\rangle z^j\left|k\right\rangle = (-1)^{jk}\left|j\right\rangle\left|k\right\rangle$.

  Somewhat unusually, this controlled gate is symmetric, where $CZ_{12} = CZ_{21}$. The matrix for this gate is

  $$CZ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

2 Quantum States as Information Carriers

Recall that information corresponds to distinguishable states of a physical system. Classically, *any* two different states are distinguishable (in principle). Quantumly, states are distinguishable reliably if and only if they are *orthogonal*. A qubit is the simplest quantum system that can reliably encode a classical bit, even though it has infinitely many different states.

So if given a quantum state $\left|\psi\right\rangle$, it is impossible to identify its identity with certainty. For example, if given $\left|0\right\rangle, \left|+\right\rangle$ as 'yes' and 'no', we would not be able to reliably distinguish between the answers.

We call 'what we get' quantum information.

Given some quantum information $\left|\psi\right\rangle$, we have three basic operations that we can perform on it.

1. *Ancilla.* We can take a fixed known quantum state $\left|A\right\rangle$ and adjoin it to get $\left|\tilde{\psi}\right\rangle = \left|\psi\right\rangle\left|A\right\rangle$. $\left|A\right\rangle$ is called an *ancilla*.

2. *Unitary.* We can apply a unitary U of our choice, with $\left|\psi\right\rangle$ becoming $U\left|\psi\right\rangle$.

3. *Measure.* We can perform a measurement on (part of) $|\psi\rangle$. We can record this result, and then use the post-measurement state for further quantum information processing. This further processing could be chosen adaptively depending on the measurement outcome.

   The output here is generally a probabilistic mixture over post-measurement states, corresponding to the Born rule.

Most generally, we can do any sequence of the three actions above.

2.1 The No-Cloning Theorem

Informally, the no-cloning theorem says that 'quantum information cannot be copied or cloned'. This is unlike the classical situation where any information can be copied.

Given some quantum information $|\psi\rangle_A$, any copy/cloning process is a quantum evolution process of A, B and M, with

$$|\psi\rangle_A |0\rangle_B |M_0\rangle_M \longrightarrow |\psi\rangle_A |\psi\rangle_B |M_\psi\rangle_M .$$

Here, B is a quantum system of the same dimension as A, usually in a 'blank', fixed state (independent of ψ), and M is any other system needed to perform the copying, stating in some 'ready'-state $|M_0\rangle$ (also independent of ψ). The final state of M may depend on $|\psi\rangle$ though.

Theorem 2.1 (No-Cloning Theorem, Unitary Version)**.** *Let $\mathcal{S}$ be any known set of quantum state of A that contains at least one pair $|\xi\rangle \neq |\eta\rangle$ of non-orthogonal states. Then no unitary process exists that achieves cloning of all states in S.*

Proof. Let $|\xi\rangle$, $|\eta\rangle$ be two different non-orthogonal states. Then the cloning process must do:

$$\begin{aligned} |\xi\rangle_A |0\rangle_B |M_0\rangle_M &\longrightarrow |\xi\rangle_A |\xi\rangle_B |M_\xi\rangle_M , \\ |\eta\rangle_A |0\rangle_B |M_0\rangle_M &\longrightarrow |\eta\rangle_A |\eta\rangle_B |M_\eta\rangle_M . \end{aligned}$$

We note that unitary operations preserve inner products, so the inner products of the LHS must equal the inner products of the RHS. So

$$\langle\xi|\eta\rangle \langle 0|0\rangle M_0 M_0 = \langle\xi|\eta\rangle \langle\xi|\eta\rangle \langle M_\xi|M_\eta\rangle .$$

Taking absolute values, we then have to have

$$|\langle\xi|\eta\rangle| = |\langle\xi|\eta\rangle|^2 |\langle M_\xi|M_\eta\rangle|.$$

Since $|\xi\rangle \neq |\eta\rangle$, and they are not orthogonal, we have $0 < |\langle\xi|\eta\rangle| < 1$, and in particular it is non-zero. But then $1 = |\langle\xi|\eta\rangle| \cdot |\langle M_\xi|M_\eta\rangle|$, which is a contradiction since $|\langle\xi|\eta\rangle| < 1$. □

Remark. Note that if $|\xi\rangle = |\eta\rangle$, then we *can* clone (trivially). Also if $|\xi\rangle$ and $|\eta\rangle$ are orthogonal, then we also can clone. We can do this by rotating to get $|0\rangle$ perpendicular to $|1\rangle$, and then applying the CX gate. Also this, this theorem is true when we include non-unitary processes (such as adding ancillas, or performing measurement).

The no cloning theorem was proved in 1982 by W. Wootters, W. Zurek, and D. Dieks independently in the same year. However, it had also been proved in 1970 by J. Park, but that went somewhat unnoticed.

This theorem was a response to N. Herbert (1980), who had a proposal for superliminal signalling in quantum mechanics. His error was to assume cloning of quantum states.

His method centered around having two participants Alice and Bob, distantly separated in space, who share an entangled state

$$\left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}(\left|0\right\rangle_A \left|0\right\rangle_B + \left|1\right\rangle_A \left|1\right\rangle_B).$$

It is easy to check that this is the same as

$$\left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}(\left|+\right\rangle_A \left|+\right\rangle_B + \left|-\right\rangle_A \left|-\right\rangle_B).$$

Alice wants to communicate a yes/no decision to Bob at noon, instantaneously. Alice decides that to communicate a 'yes', she will measure in the $\{\left|0\right\rangle, \left|1\right\rangle\}$ basis, and to communicate a 'no', she will measure in the $\{\left|+\right\rangle, \left|-\right\rangle\}$ basis.

The Born rule implies that Bob's qubit is instantaneously collapsed. So for the 'yes' action, with 50/50 probability he will be holding $|0\rangle$ or $|1\rangle$. For the no action, with 50/50 probability he will be holding $|+\rangle$ or $|-\rangle$.

But these yes/no preparations of $B's$ qubits are *indistinguishable* by any local action for Bob. For any measurement, the output probabilities are the same in both cases.

Indeed, if Π_i is the projection operator for outcome i in a measurement for Bob, then in the yes case, $p_{\text{yes}}(i) = \frac{1}{2}\langle 0|\Pi_i|0\rangle + \frac{1}{2}\langle 1|\Pi_i|1\rangle = \text{tr}\left(\Pi_i\left(\frac{|0\rangle\langle 0|+|1\rangle\langle 1|}{2}\right)\right)$. In the no case, $p_{\text{no}}(i) = \text{tr}\left(\Pi_i\left(\frac{|+\rangle\langle +|+|-\rangle\langle -|}{2}\right)\right)$, but $|0\rangle\langle 0| + |1\rangle\langle 1| = |+\rangle\langle +| + |-\rangle\langle -| = I$, and the probabilities are equal. Thus B cannot detect $A's$ attempted signalling.

Now if B could *clone* quantum states, then just after noon, Bob could clone his qubit many times (say 10^6). Now in the 'yes' case, all will be zero or all will be one. Similarly in the 'no' case, he measures all copies in the $\{\left|0\right\rangle, \left|1\right\rangle\}$ basis and he will see a uniformly random bit string. Of course this is not possible, by the above result.

3 Distinguishing Non-Orthogonal States

We know that we can't perfectly distinguish non-orthogonal states. But how well can we do it?

- *Given.* An unknown quantum state $|\psi\rangle$ of dimension d.
- *Promise.* $|\psi\rangle$ is either $|\alpha_0\rangle$ or $|\alpha_1\rangle$, which are distinct known states.
- *Problem.* Determine which i is such that $|\psi\rangle = |\alpha_i\rangle$.

We will apply the *state estimation* process.

Given $|\psi\rangle$, we can adjoin an ancilla $|A\rangle$ to give $|\psi\rangle\,|A\rangle$. Then we can perform a unitary on the full system, and we can perform a measurement with outcomes 0 and 1 for our answer.

Remark. It can be shown (as with the no-cloning theorem) that this process is completely general.

We can simplify this process. Adjoining $|A\rangle$ just changes discrimination of $|\alpha_0\rangle$ vs $|\alpha_1\rangle$ to that of $|\alpha_0\rangle\,|A\rangle$ vs $|\alpha_1\rangle\,|A\rangle$, with the same inner product. So this is just working in a larger dimension (and we will see later that the dimension will not matter).

If we call the process M_1 'doing U followed my measurement with projectors Π_0, Π_1', it is equivalent to just performing M_2 (the 'U-rotated measurement') with projectors $\tilde{\Pi}_i = U^\dagger \Pi_i U$ as they give the same outputs on any state $|\xi\rangle$,

$$\text{prob}_{M_1}(i) = (\langle\xi|\,U^\dagger)\Pi_i(U\,|\xi\rangle) = \text{prob}_{M_2}(i).$$

Hence our process is equivalent to just a single measurement.

Given $|\alpha_0\rangle$ or $|\alpha_1\rangle$, we can just perform a single measurement with projectors Π_0 and Π_1. Of course, some measurements are better than others for querying the correct answer with higher probability. We will introduce a 'figure of merit' for the measurement, the *success probability*.

Definition 3.1 (Success Probability)**.** With no prior knowledge of which state we are getting, we assume the prior probabilities of a half. Then the *success probability* P_S is

$$P_s = \frac{1}{2}\,\text{prob}(\text{output is } 0 \mid |\alpha_0\rangle \text{ input}) + \frac{1}{2}\,\text{prob}(\text{output is } 1 \mid |\alpha_1\rangle \text{ input}).$$

so by the Born rule,

$$P_s = \frac{1}{2}(\,\langle\alpha_0|\Pi_0|\alpha_0\rangle + \langle\alpha_1|\Pi_1|\alpha_1\rangle).$$

Since $\Pi_0 + \Pi_1 = I$, $\Pi_1 = I - \Pi_0$, so

$$\begin{aligned} P_S &= \frac{1}{2} + \frac{1}{2}(\,\langle\alpha_0|\Pi_0|\alpha_0\rangle - \langle\alpha_1|\Pi_0|\alpha_1\rangle) \\ &= \frac{1}{2} + \frac{1}{2}\,\text{tr}\left[\Pi_0(|\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|)\right], \end{aligned}$$

and we can calculate the optimal $\{\Pi_0, I - \Pi_0\}$. We look more closely at

$$D = |\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|\,.$$

This has the following properties:

- D is hermitian, so it has real eigenvalues and has a complete basis of orthonormal eigenstates.
- If $|\beta\rangle \perp |\alpha_0\rangle$ and $|\alpha_1\rangle$, then $D\,|beta\rangle = 0$. So D has only two non-zero eigenvalues, and the eigenvectors are spanned by $|\alpha_0\rangle$ and $|\alpha_1\rangle$.

- $\operatorname{tr} D = 0$, so the non-zero eigenvalues must sum to 0. Thus we can write them as $+\delta$ and $-\delta$, with $\delta > 0$ and eigenstates $|p\rangle$ and $|m\rangle$.
- We can write
$$D = \delta\,|p\rangle\langle p| - \delta\,|m\rangle\langle m|\,.$$
Now we need to determine δ in terms of $|\alpha_0\rangle$ and $|\alpha_1\rangle$.
- We can work in the two dimensional subspace of $|\alpha_0\rangle$ and $|\alpha_1\rangle$. We will introduce an orthonormal basis so that we can work with components. We choose $\left|\alpha_0^\perp\right\rangle$ orthonormal to $|\alpha_0\rangle$, and we use $\{\left|\alpha_0^\perp\right\rangle, |\alpha_0\rangle\}$ as a basis. Then
$$|\alpha_0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad |\alpha_1\rangle = \begin{pmatrix}c_0\\c_1\end{pmatrix} = c_0\,|\alpha_0\rangle + c_1\left|\alpha_0^\perp\right\rangle.$$
So $c_0 = \langle\alpha_0|\alpha_1\rangle$, and $|c_1| = \sin\theta$ for $\cos\theta = |\,\langle\alpha_0|\alpha_1\rangle\,|$.
- We can then write down the projectors
$$D = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} - \begin{pmatrix}c_0\\c_1\end{pmatrix}\begin{pmatrix}c_0^* & c_1^*\end{pmatrix} = \begin{pmatrix}|c_1|^2 & -c_0c_1^*\\ -c_1c_0^* & -|c_1|^2\end{pmatrix},$$
then $\det(D - \lambda I) = 0$ gives $\lambda = \pm|c_1| = \pm\sin\theta$. So $\delta = \sin\theta$.

Finally returning to our expression for the success probability, we can use our new expression for D to get

$$\begin{aligned} P_s &= \frac{1}{2} + \frac{\delta}{2}\operatorname{tr}\left[\Pi_0(|p\rangle\langle p| - |m\rangle\langle m|)\right] \\ &= \frac{1}{2} + \frac{\delta}{2}\left(\,\langle p|\Pi_0|p\rangle - \,\langle m|\Pi_0|m\rangle\right). \end{aligned}$$

Now for any projector Π and state $|\xi\rangle$, we have $0 \le \langle\xi|\Pi|\xi\rangle \le 1$. Hence P_S achieves its maximum value of $\frac{1}{2} + \frac{\delta}{2} = \frac{1}{2}(1 + \sin\theta)$ if Π_0 is chosen to be any projector into a subspace containing $|p\rangle$ so $\Pi_0\,|p\rangle = |p\rangle$ and is orthogonal to $|m\rangle$, so that $\Pi_0\,|m\rangle = |0\rangle$. Such a choice of Π_0 is always possible since $|p\rangle$ is orthogonal to $|m\rangle$. The achievable bound is then

$$P_S \le \frac{1}{2} + \frac{\sin\theta}{2},$$

which is the *Helstrom-Holevo bound* (for pure states).

In particular, an ancilla is never needed, only the inner product is relevant. Also, if $|\alpha_0\rangle$, $|\alpha_1\rangle$ are qubit states (with dimension 2), we can work entirely in their two dimensional space, and the optimal discriminating measurement will be a complete projective measurement of the quantum observable $D = |\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|$, that is, measurement relative to the eigenbasis of D.

We have thus proved the following theorem.

Theorem 3.2 (Helstrom-Holevo Bound)**.** *Given one of two equally likely states, α_0 and α_1 with $|\,\langle\alpha_0|\alpha_1\rangle\,| = \cos\theta$, the probability P_S of correctly identifying the state by any quantum process is bounded by $P_S \le \frac{1}{2}(1 + \sin\theta)$, and the bound is tight.*