Part III **January 2014**

QUANTUM COMPUTATION PREREQUISITE MATERIAL

Richard Jozsa rj310@cam.ac.uk

A basic knowledge of linear algebra, Dirac bra-ket notation and the postulates of quantum mechanics will be assumed as prerequisites for this course. The required material (and a bit more!) is contained in the sections below. Some of this material will be further explained in lectures but you should have at least a prior acquaintance with it e.g. by reading through these notes, even if you don't initially understand everything.

These notes are meant to be suitable for assimilating the material if unfamiliar. For further explanations see:
M. Nielsen and I. Chuang "Quantum computation and information". CUP.
John Preskill's notes for Caltech course on quantum computation available at
http://www.theory.caltech.edu/ preskill/ph219/index.html♯lecture

At the end we include some exercises, which in addition to illustrating the formalism, are of interest in themselves (and recommended even if you are already familiar with the prerequisite material).

1 Linear algebra, Dirac notation

Bra and ket vectors

Let V be a (finite dimensional) complex vector space of dimension m with an inner product. Vectors in V will be written as $|v\rangle$ and called *ket vectors* or just *kets*. In this course we'll often work with a 2-dimensional space V_2 with a chosen orthonormal basis denoted $\{|0\rangle, |1\rangle\}$. (All constructions and formulae can be easily generalised to arbitrary finite dimensional spaces). Ket vectors $|v\rangle = a\,|0\rangle + b\,|1\rangle$are always written in components as *column* vectors

$$|v\rangle = \begin{pmatrix} a \\ b \end{pmatrix}.$$

The conjugate transpose $|v\rangle^\dagger$ (denoted by a dagger) is called a *bra vector* and is written using a mirror image bracket

$$\langle v| = |v\rangle^\dagger = a^*\,\langle 0| + b^*\,\langle 1| = (a^*\ b^*).$$

Thus in components bra vectors are always written as row vectors. If $|w\rangle = c\,|0\rangle + d\,|1\rangle$ is another ket then the inner product of $|v\rangle$ and $|w\rangle$ is written by juxtaposing brackets

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

(Indeed the whole Dirac notation formalism is motivated by the bracket notation $(\underline{v}, \underline{w})$ for inner products commonly used in mathematics.) Orthonormality of the basis $\{|0\rangle, |1\rangle\}$ is equivalent to the condition $\langle i|j\rangle = \delta_{ij}$ (the Kronecker delta). In more abstract terms, bra vectors are a notation for elements of the dual space V^* viz. $\langle v|$ is the linear functional whose value on any ket $|w\rangle$ is the inner product $\langle v|w\rangle$.

Tensor products of vectors

If V and W are vector spaces of dimensions m and n with bases $\{|e_1\rangle, \ldots, |e_m\rangle\}$ and $\{|f_1\rangle, \ldots, |f_n\rangle\}$

respectively, then the tensor product space $V \otimes W$ has dimension mn and can be regarded as consisting of all formal linear combinations of the symbols $|e_i\rangle \otimes |f_j\rangle$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. There is a natural bilinear embedding $V \times W \to V \otimes W$ defined as follows. If $|\alpha\rangle = \sum_i a_i |e_i\rangle$ and $|\beta\rangle = \sum_j b_j |f_j\rangle$ are general vectors in V and W respectively then

$$(|\alpha\rangle, |\beta\rangle) \mapsto |\alpha\rangle \otimes |\beta\rangle = \sum_{ij} a_i b_j |e_i\rangle \otimes |f_j\rangle$$

obtained by formally "multiplying out" the brackets in $(\sum_i a_i |e_i\rangle)(\sum_j b_j |f_j\rangle)$. Any such vector $|\alpha\rangle \otimes |\beta\rangle$ is called a *product vector*. The mapping is not surjective – vectors in $V \otimes W$ that are not product vectors are called *entangled vectors*. We often write the product vector $|\alpha\rangle \otimes |\beta\rangle$ simply as $|\alpha\rangle |\beta\rangle$ (omitting the $\otimes$).

The inner products on V and W give a natural inner product on $V \otimes W$ defined "slot-wise" (the slots being the component spaces). Thus for product vectors we have the inner product of $|\alpha_1\rangle |\beta_1\rangle$ with $|\alpha_2\rangle |\beta_2\rangle$ being $\langle\alpha_1|\alpha_2\rangle\langle\beta_1|\beta_2\rangle$. This extends to general (entangled) vectors by linearity since general vectors are always linear combinations of product vectors (e.g. the product basis vectors $|e_i\rangle |f_j\rangle$).

Note that we generally write the bra vector of $|\alpha\rangle |\beta\rangle \in V \otimes W$ as $\langle\beta| \langle\alpha|$ with order of spaces reversed. If we need to make the associated spaces explicit they can be denoted by subscripts e.g. $|\alpha\rangle_V |\beta\rangle_W$ and ${}_W\langle\beta| {}_V\langle\alpha|$.

We will be mostly concerned with tensor products of the two-dimensional space V_2 with itself (multiple times). For the k-fold tensor power we write $\otimes^k V_2 = V_2 \otimes \ldots \otimes V_2$ which is a space of dimension 2^k with basis $|i_1\rangle \otimes \ldots \otimes |i_k\rangle$ $(i_1, \ldots i_k = 0, 1)$ labelled by the 2^k k-bit strings $i_1 \ldots i_k$. We often write $|i_1\rangle \otimes \ldots \otimes |i_k\rangle$ simply as $|i_1 \ldots i_k\rangle$.

Example. For $k = 2$ if $|v\rangle = a|0\rangle + b|1\rangle$ and $|w\rangle = c|0\rangle + d|1\rangle$, we have $|v\rangle |w\rangle \in V_2 \otimes V_2$ and in terms of components we have

$$|v\rangle \otimes |w\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \begin{pmatrix} c \\ d \end{pmatrix} \\ b \begin{pmatrix} c \\ d \end{pmatrix} \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}.$$

Furthermore an arbitrary vector $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle \in V_2 \otimes V_2$ is entangled iff $ad - bc \neq 0$ (exercise). $\square$

Dirac notation for linear maps

With $|v\rangle$ and $|w\rangle$ in V_2 as above, standard matrix multiplication gives

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

which is a linear map on V (acting by matrix multiplication on column vectors). In fact for any $|x\rangle \in V_2$ we have $M|x\rangle = |v\rangle \langle w|x\rangle$, a vector in the direction of $|v\rangle$. In particular if $|v\rangle$ is normalised (i.e. $\langle v|v\rangle = |a|^2 + |b|^2 = 1$) then $\Pi_v = |v\rangle\langle v|$ is the operator of *projection onto* $|v\rangle$, satisfying $\Pi_v \Pi_v = \Pi_v$. The latter can be seen very neatly in Dirac notation: $\Pi_v \Pi_v = (|v\rangle\langle v|)(|v\rangle\langle v|) = |v\rangle \langle v|v\rangle \langle v| = |v\rangle\langle v| = \Pi_v$ as $\langle v|v\rangle = 1$.

Note that

$$|0\rangle\langle 0| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad |0\rangle \langle 1| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{etc.}$$

so if $A : V_2 \to V_2$ is any linear map with matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then we can write

$$A = a\,|0\rangle\langle 0| + b\,|0\rangle\langle 1| + c\,|1\rangle\langle 0| + d\,|1\rangle\langle 1|\,.$$

Note that if $|x\rangle = A\,|v\rangle$ then $\langle x| = (A\,|v\rangle)^\dagger = |v\rangle^\dagger\,A^\dagger = \langle v|\,A^\dagger$.

Tensor products of maps
If

$$B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

is a second linear map on V_2 then the tensor product of maps $A \otimes B : V_2 \otimes V_2 \to V_2 \otimes V_2$ is defined by its action on the basis $|i\rangle\,|j\rangle \to A\,|i\rangle\,B\,|j\rangle$ for i and j being 0,1, or more generally on product vectors by $(A \otimes B)(|v\rangle\,|w\rangle) = A\,|v\rangle \otimes B\,|w\rangle$. It has a simple 4×4 matrix of components, with block form

$$A \otimes B = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix} = \begin{pmatrix} ap & aq & bp & bq \\ ar & as & br & bs \\ cp & cq & dp & dq \\ cr & cs & dr & ds \end{pmatrix}.$$

In particular we have $A \otimes I$ and $I \otimes A$ being the action of A on the first (resp. second) component space of $V_2 \otimes V_2$.

Example: for $|\psi\rangle = |00\rangle + |11\rangle$ and A as above, we have $A \otimes I\,|\psi\rangle = (A\,|0\rangle)\,|0\rangle + (A\,|1\rangle)\,|1\rangle = (a\,|0\rangle + c\,|1\rangle)\,|0\rangle + (b\,|0\rangle + d\,|1\rangle)\,|1\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle$. $\square$

Partial inner products for vectors in $V \otimes W$
It is useful to introduce the following "partial" operations on component spaces in tensor products. Any $|v\rangle \in V$ defines a linear map $V \otimes W \to W$ called "*partial inner product* with $|v\rangle$". It is defined on the basis of $V \otimes W$ by the formula $|e_i\rangle\,|f_j\rangle \to \langle v|e_i\rangle\,|f_j\rangle$. Similarly any $|w\rangle \in W$ defines a map $V \otimes W \to V$.

Example. For $|v\rangle \in V$ and $|\xi\rangle \in V \otimes V$ we can form the partial inner product on the first or second space. To make the position explicit we sometimes introduce subscripts to label the slots, writing $V \otimes V$ as $V_{(1)} \otimes V_{(2)}$, and writing ${}_1\langle v|\xi\rangle_{12} \in V_{(2)}$ and ${}_2\langle v|\xi\rangle_{12} \in V_{(1)}$ for the partial inner products. Thus for example, if $|\xi\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle$ then the orthonormality relations $\langle i|j\rangle = \delta_{ij}$ give ${}_1\langle 0|\xi\rangle_{12} = a\,|0\rangle + b\,|1\rangle$ and ${}_2\langle 0|\xi\rangle_{12} = a\,|0\rangle + c\,|1\rangle$ i.e. we just pick out the terms of $|\xi\rangle$ that contain 0 in the first, respectively second, slot.

Partial traces of linear maps on $V \otimes W$ (optional)
If $M : V \otimes W \to V \otimes W$ is any linear operation on a tensor product space, then the *partial traces* over V or W are respectively the linear maps on W and V denoted $\mathrm{Tr}_V M : W \to W$ and $\mathrm{Tr}_W M : V \to V$, defined by

$$\mathrm{Tr}_V M = \sum_{i=1}^{m} {}_1\langle e_i|M|e_i\rangle_1 \qquad \mathrm{Tr}_W M = \sum_{j=1}^{n} {}_2\langle f_j|M|f_j\rangle_2$$

i.e. we take partial inner products with an orthonormal basis and sum over the basis elements, to "trace out" that space. This construction is independent (exercise) of choice of orthonormal basis in V or W. Thus we can express the full trace as a sequence of partial traces $\mathrm{Tr}\,M = \mathrm{Tr}_W(\mathrm{Tr}_V M)$.

Fact. (exercise) If $M = A \otimes B$ on $V_2 \otimes V_2$ then $\mathrm{Tr}_1 M = (\mathrm{Tr}\,A)\,B$ and $\mathrm{Tr}_2 M = A\,(\mathrm{Tr}\,B)$. For any M on $V_2 \otimes V_2$ with block form (and written with respect to the standard basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$)

$$M = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

(with P, Q, R, S being 2×2 sized submatrices) we have

$$\mathrm{Tr}_1 M = P + S \qquad \mathrm{Tr}_2 M = \begin{pmatrix} \mathrm{Tr}\, P & \mathrm{Tr}\, Q \\ \mathrm{Tr}\, R & \mathrm{Tr}\, S \end{pmatrix}$$

and similarly for general M on any $V \otimes W$ with corresponding block form ($m \times m$ blocks of $n \times n$ sized sub-matrices, where m and n are the dimensions of V and W respectively.)

2 Principles of quantum mechanics (QM1) – (QM4)

Our description of quantum mechanics below may at first sight look a little different from standard textbook presentations but in fact it's equivalent. Here we focus on finite-dimensional quantum mechanics (multi-qubit systems, unitary matrices representing finite time evolutions etc.) whereas physics textbooks usually begin with the infinite dimensional case (wavefunctions, Schrödinger's wave equation giving infinitesimal time evolution via a Hamiltonian etc.) and we also emphasise *ab initio* the quantum measurement formalism which will be of crucial significance for us.

(QM1) (physical states): the states of any (isolated) physical system are represented by unit vectors in a complex vector space with an inner product (but see also the remark about global and relative phases at the end of this section). Two states are physically distinguishable iff the corresponding vectors are orthogonal. □

We emphasise here that in classical physics, *any* two different states of a system are in principle distinguishable, but in quantum theory this is no longer the case. This important novel feature will be quantified in (QM4) (quantum measurements) below.

The simplest non-trivial quantum system has states lying in a 2 dimensional vector space, thus allowing only two mutually distinguishable states. Choosing a pair of orthonormal vectors and labelling them $|0\rangle$ and $|1\rangle$, the general state can be written

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle \qquad |a|^2 + |b|^2 = 1.$$

We say that $|\psi\rangle$ is a *superposition* of states $|0\rangle$ and $|1\rangle$ with *amplitudes* a and b.

Qubits: any quantum system, with a 2 dimensional state space and with a chosen orthonormal basis $\{|0\rangle, |1\rangle\}$ is called a *qubit*. The basis states $|0\rangle, |1\rangle$ are called *computational basis states* or *standard* basis states. There are many real physical systems that can embody the structure of a qubit, for example the spin of an electron, the polarisation of a photon, superpositions of two selected energy levels in an atom etc.

(QM2) (composite systems): if system S_1 had state space V and system S_2 has state space W then the joint system obtained by taking S_1 and S_2 together, has states given by arbitrary unit vectors in the *tensor product* space $V \otimes W$. □

Basic example: (n qubits) thus a system comprising n qubits has a state space of dimension 2^n An n-qubit state $|\psi\rangle$ is a called a *product state* if it is the product of n single-qubit states $|\psi\rangle = |v_1\rangle\,|v_2\rangle \ldots |v_n\rangle$ and $|\psi\rangle$ is called *entangled* if it is not a product state.

We note the significant fact that as the number of qubits grows *linearly*, the full state description (given as the full list of amplitudes) grows *exponentially* in its complexity. However the description of any product state grows only linearly with n (each successive $|v_i\rangle$ is described by two further amplitudes) so this exponential complexity of state description is intimately related to the phenomenon of entanglement that arises for tensor products of spaces. With this in

mind, it is especially interesting to contrast (QM2) with its classical counterpart – for *classical* physics, the state space of a composite system is the *cartesian* product of the state spaces of the constituent parts. Thus if classical system S requires K parameters for its state description then a composite of n such systems will require only nK parameters i.e. a linear growth of description, in contrast to the exponential growth for quantum systems. □

(QM3) (physical evolution of quantum systems): any physical (finite time) evolution of an (isolated) quantum system is represented by a *unitary* operation on the corresponding vector space of states. □

Recall that a linear operation U on any vector space is unitary if its matrix has $U^{-1} = U^\dagger$ (where dagger is conjugate transpose) or equivalently, if the rows (or columns) of the matrix U form an orthonormal basis of vectors.

(QM4) (quantum measurements, the Born rule). In classical physics the state of any given physical system can always in principle be fully determined by suitable "measurements" on a single instance of the system, while leaving the original state intact. In quantum theory the corresponding situation is bizarrely different – quantum measurements generally have only probabilistic outcomes, they are "invasive", generally unavoidably destroying the input state and they reveal only a rather small amount of information about the (now irrevocably lost) input state identity. Furthermore the (probabilistic) change of state in a quantum mesaurement is (unlike normal time evolution) not a unitary process. Here we outline the associated mathematical formalism, which is at least, easy to apply.

The basic Born rule: suppose we are given a (single physical instance of a) quantum state for a system with state space V of dimension n. Let $\mathcal{B} = \{|e_1\rangle, \ldots, |e_n\rangle\}$ be any orthonormal basis of V and write $|\psi\rangle = \sum a_i |e_i\rangle$. Then we can make a *quantum measurement of* $|\psi\rangle$ *relative to the basis* $\mathcal{B}$. (In textbooks this is often called a (complete) von Neumann measurement). The possible outcomes are $j = 1, \ldots, n$ corresponding to the basis states $|e_j\rangle$. The probability of obtaining outcome j is

$$\mathrm{pr}(j) = |\langle e_j|\psi\rangle|^2 = |a_j|^2.$$

If outcome j is seen then after the measurement the state is no longer $|\psi\rangle$ but has been "collapsed" to $|\psi_{\mathrm{after}}\rangle = |e_j\rangle$ i.e. the basis state corresponding to the seen outcome. Stated alternatively: the probability is the squared length of the projection of $|\psi\rangle$ onto the basis state, and the post-measurement state is that projected vector, renormalised to have length 1.
Remark. In textbooks we often read of measurement of a *quantum observable* which is just a slight variation of the above. A quantum observable is defined to be a Hermitian operator A on V. Recall that a Hermitian matrix always has real eigenvalues λ_i (for simplicity taken to be non-degenerate) and there is an orthonormal basis of associated eigenvectors, denoted $|\lambda_i\rangle$ i.e. the observable encodes a basis, and the measurement of the observable is just a measurement in our sense above relative to this basis (with probabilities and post-measurement states as above), but with outcomes labelled by the eigenvalues λ_j rather than the labels j themselves. □
The extended Born rule. In this course we will often consider measurement of only some part of a composite system, and the associated formalism for probabilities, outcomes and post-measurement states is called the extended Born rule. Suppose $|\psi\rangle$ is a quantum state of a composite system S_1S_2 with state space $V \otimes W$. Let $\mathcal{B} = \{|e_1\rangle, \ldots, |e_n\rangle\}$ be an orthonormal basis of V. Note that $|\psi\rangle$ can be expanded uniquely as $|\psi\rangle = \sum_i |e_i\rangle |\xi_i\rangle$ with $|\xi_i\rangle$ being vectors in W (not generally normalised nor orthogonal). Indeed orthonormality of the basis gives the $|\xi_k\rangle$'s as the partial inner products $|\xi_k\rangle = \langle e_k|\psi\rangle$. Alternatively if $\{|f_1\rangle, \ldots, |f_m\rangle\}$ is a basis of W then writing $|\psi\rangle = \sum_{ij} a_{ij} |e_i\rangle |f_j\rangle$ in the product basis of $V \otimes W$, we see that $|\xi_k\rangle = \sum_j a_{kj} |f_j\rangle$ i.e. we just pick out all terms of $|\psi\rangle$ that involve $|e_k\rangle$ in the V-slot.

Now we can make a measurement of $|\psi\rangle \in V \otimes W$ relative to the basis $\mathcal{B}$ of V. The extended Born rule asserts the following:
(a) the probability of outcome $k = 1, \ldots, n$ is $\mathrm{pr}(k) = \langle \xi_k | \xi_k \rangle$ i.e. the squared length of the partial inner product $\langle e_k | \psi \rangle$;
(b) if the outcome k is seen then the post-measurement state is the product state

$$|\psi_{\text{after}}\rangle = |e_k\rangle \, |\xi_k\rangle \, / \sqrt{\mathrm{pr}(k)}$$

i.e. the V-slot is "collapsed" to the seen outcome $|e_k\rangle$ and the W-slot retains only the associated vector $|\xi_k\rangle$ but renormalised by $\sqrt{\mathrm{pr}(k)}$ to have unit length.

Note that the basic Born rule is just a special case of (a) and (b) with W having dimension 1 and $|\xi_k\rangle = \langle e_k | \psi \rangle$ is just the complex number a_k (viewed as a 1-dimensional vector).

Fixed choice of basis: note that a measurement relative to any general basis $\mathcal{C}$ can be performed by a measurement relative to any a priori fixed basis $\mathcal{B}$ together with some unitary operations; indeed any two orthonormal bases $\mathcal{B} = \{|e_1\rangle, \ldots, |e_n\rangle\}$ and $\mathcal{C} = \{|e'_1\rangle, \ldots, |e'_n\rangle\}$ are related by a unitary transformation U as $|e'_i\rangle = U \, |e_i\rangle$. Thus to perform a measurement on $|\psi\rangle$ relative to $\mathcal{C}$ we first apply U^{-1} to $|\psi\rangle$, then perform a measurement relative to $\mathcal{B}$, then finally apply U to the resulting post-measurement state.
Standard measurement on multi-qubit systems. Recall that any k-qubit system comes equipped with a standard or computational basis $\mathcal{B}$ of orthonormal states labelled by k-bit strings. In this course our measurements will often be restricted to being only relative to this standard basis for some subset of k qubits of an n-qubit system.
Example. Consider the 3-qubit state

$$|\phi\rangle = \frac{i}{2} \, |000\rangle + \frac{12 + 5i}{26} \, |001\rangle - \frac{1}{2} \, |101\rangle + \frac{3}{10} \, |110\rangle - \frac{2i}{5} \, |111\rangle \, .$$

Computing the partial inner product with $|1\rangle$ on the first qubit we get

$$|\alpha\rangle = {}_1\langle 1 | \phi \rangle = -\frac{1}{2} \, |01\rangle + \frac{3}{10} \, |10\rangle - \frac{2i}{5} \, |11\rangle$$

and its squared length is $\langle \alpha | \alpha \rangle = 1/2$. Hence if we make a standard measurement on the first qubit, the probability of seeing outcome 1 is half, and in that case the state after the measurement will be $|1\rangle$ "$|\alpha\rangle$ normalised" $= \sqrt{2} \, |1\rangle \, |\alpha\rangle$. $\square$

Remark (global and relative phases) If $|v\rangle$ is any unit vector then the states $|v\rangle$ and $e^{i\alpha} \, |v\rangle$ will have the same measurement probabilities (for any basis), independent of α (since probabilities always depend on squared moduli of amplitudes.) Here α is called a *global phase*. Thus $|v\rangle$ and $e^{i\alpha} \, |v\rangle$ represent identical physical situations and in (QM1) we should (more correctly) have said that states of a physical system correspond to unit vectors *up to an (irrelevant) global phase*. Note also that the projection operator $\Pi_v = |v\rangle\langle v|$ is independent of the choice of global phase for $|v\rangle$ and hence it can also be used to uniquely represent distinct physical systems.
On the other hand θ in $|0\rangle + e^{i\theta} \, |1\rangle$ is called a *relative* (or local) phase and it is a crucially important parameter for the qubit state. Indeed for example, we can think of any unitary operation as evolving $|0\rangle$ and $|1\rangle$ separately and combining the results with relative phase θ which will affect the way that the two terms interfere. A notable example is the Hadamard operation H (cf below) acting on the states $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. These differ only by a relative phase π but we have $H \, |+\rangle = |0\rangle$ and $H \, |-\rangle = |1\rangle$. $\square$

3 Some basic unitary operations for qubits

Unitary operations on qubits are also called *quantum gates*. Matrices given below are always relative to the standard basis $|0\rangle, |1\rangle$.

One-qubit gates

$$\textbf{Hadamard gate} \qquad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Introduce the 1-qubit gates

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad Y = ZX = -XZ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

(X is sometimes also called the NOT-gate).
Then the **Pauli operations** are

$$\sigma_x = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = -iY = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

They have elegantly simple multiplicative properties:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$$
$$\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z \qquad \sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x \qquad \sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y$$

(noting the cyclic shift of x, y, z labels in the latter set). Note that the matrices $I, \sigma_x, \sigma_y, \sigma_z$ are all Hermitian as well as unitary (which is an unusual coincidence).

$$\textbf{Phase gate} \qquad P_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}.$$

Two-qubit gates

Controlled-X (or controlled-NOT) gate

$$CX = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} (I) & (0) \\ (0) & (X) \end{pmatrix}.$$

For the four basis states we can compactly write $CX\,|i\rangle\,|j\rangle = |i\rangle\,|i \oplus j\rangle$ where $\oplus$ denotes addition modulo 2. Note that for any 1-qubit state $|\alpha\rangle$ we have

$$CX\,|0\rangle\,|\alpha\rangle = |0\rangle\,|\alpha\rangle \qquad CX\,|1\rangle\,|\alpha\rangle = |1\rangle\,X\,|\alpha\rangle$$

i.e CX applies X to the second qubit if the first is set to "1" and acts as the identity if the first is set to "0" (and extends by linearity if the first qubit is in a superposition of the two values etc.) Accordingly the first qubit is called the *control qubit* and the second is called the *target qubit*.

The **controlled-Z gate**:

$$CZ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} (I) & (0) \\ (0) & (Z) \end{pmatrix}$$

i.e. as for CX, CZ applies Z to the second qubit controlled by the state of the first qubit. Note that despite this asymmetrical description, CZ (unlike CX) is actually symmetric in its action on the two qubits.

4 Exercises

(1) (Quantum teleportation) Write $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and let $|\alpha\rangle = a\,|0\rangle + b\,|1\rangle$ be a general 1-qubit state. Subscripts will denote qubit positions labelled from left to right, in a multi-qubit state.
(a) Write $|A\rangle_{123} = |\alpha\rangle_1 |\psi\rangle_{23}$ in the computational basis of three qubits and hence compute $|B\rangle_{123} = (H_1 \otimes I_{23})(CX_{12} \otimes I_3)\,|A\rangle_{123}$.
(b) Suppose we perform a standard quantum measurement on qubits 1 and 2 of $|B\rangle$. Show that the four possible outcomes $ij = 00, 01, 10, 11$ are always equiprobable and compute the post-measurement state in each case.
(c) Show that in each case the post-measurement state in slot 3 is a unitary transform of $|\alpha\rangle$ (independent of a and b) and identify the corresponding unitary matrix U_{ij} for each possible outcome ij.
Remark: in quantum teleportation Alice holds qubits 1 and 2 while Bob, distantly separated in space, holds qubit 3. So Alice, by applying the local operations H_1, CX_{12} and local measurements, can faithfully transfer the state of qubit 1 to Bob (even if she does not know its identity), at the communication expense of sending him only *two classical bits* ij (so he can correct the unitary "error" U_{ij}).

(2) (Basic entanglement) Prove that the state $a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle$ is entangled iff $ad - bc \neq 0$. Deduce that the state $|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + (-1)^k\,|11\rangle)$ is entangled if $k = 1$ and unentangled if $k = 0$. Express the latter case explicitly as a product state. How can $|\psi\rangle$ (for $k = 0, 1$) be manufactured starting from $|00\rangle$ and applying only gates from those listed in section 3 above?

(3) (No cloning of quantum states) We routinely copy classical data in everyday life e.g. for a single bit value $b = 0$ or 1, show that the classical CNOT gate (which operates just like the quantum CX gate on basis states viz. $(b, c) \mapsto (b, b \oplus c)$ for bits b, c) when acting on the 2-bit pair $(b, 0)$, will copy b into the second slot i.e. we get (b, b).
(i) Consider now the quantum CNOT gate acting on the 2-qubit state $|\psi\rangle\,|0\rangle$ where $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ is a general qubit state. Will we now get a copy of $|\psi\rangle$ in the second register? i.e. do we get $|\psi\rangle\,|\psi\rangle$?
(ii) Consider *any* process which purports to clone an arbitrary input qubit state. Any such process has the following form. The input is $|\psi\rangle\,|0\rangle \ldots |0\rangle$ where $|\psi\rangle$ is any qubit state and $|0\rangle \ldots |0\rangle$ are any required number of "working space" qubits all in state $|0\rangle$. The output is $|\psi\rangle\,|\psi\rangle\,|A_\psi\rangle$ i.e. we get two copies of $|\psi\rangle$ together with (possibly) some further ψ-dependent state $|A_\psi\rangle$. Prove that no such process can exist within the framework of quantum theory i.e. "quantum states cannot be cloned". (Hint: think about unitarity).

(4) (Quantum nonlocality) Consider the 2-qubit state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. We imagine that the two qubits are separated at great spatial distance and held by Alice (A) and Bob (B) respectively, who can then apply quantum operations (unitary gates and measurements) only to the qubit they hold. Introduce the 1-qubit gate ("rotation by θ")

$$U(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

(a) Suppose A applies $U(\alpha)$ and B applies $U(\beta)$. Show that the resulting state is

$$|\psi_{\alpha\beta}\rangle = \tfrac{1}{\sqrt{2}}\left(\cos(\alpha-\beta)\,|00\rangle - \sin(\alpha-\beta)\,|01\rangle + \sin(\alpha-\beta)\,|10\rangle + \cos(\alpha-\beta)\,|11\rangle\right).$$

Deduce that for any choice of α and β, if we measure either one of the qubits of $|\psi_{\alpha\beta}\rangle$ in the computational basis we will get output 0 or 1 with equal probabilities of half. Show that this remains true even if the other party has (unbeknown to us) already made the measurement

and obtained his/her suitably random outcome and post-measurement state i.e. choice of local angle on one side cannot affect the measurement statistics obtained locally on the other side.
(b) Suppose A and B now both measure their held qubit of $|\psi_{\alpha\beta}\rangle$ (in either order and pooling their results, or simultaneously as a 2-qubit measurement – the statistics are the same). Show that pr(outcomes differ) $= \sin^2(\alpha - \beta)$. (Note that "outcomes differ" means "outcome is 01 or 10").
(c) We now consider only three angle settings $\theta = -\frac{\pi}{6}, 0, \frac{\pi}{6}$. Let $M_A(\alpha)$ denote the following operation for Alice: apply $U(\alpha)$ to her qubit and measure it in the computational basis. Similarly $M_B(\beta)$ for Bob. Consider now the following experiment denoted $E(\alpha, \beta)$: A and B have many $|\psi\rangle$ states and perform a long sequence of $M_A(\alpha)$ and $M_B(\beta)$ with each choosing one of the allowed angles (which is kept the same for the whole sequence). We imagine that for each $|\psi\rangle$ the local operations are done essentially simultaneously (or at least at a spacelike interval). For long sequences, probabilities will be reflected in frequencies of occurrence of 0's and 1's. Show that the following statistics will be seen:

(i) $E(0,0)$:	pr(differ) $= 0$	A and B's sequences will be the same sequence.
(ii) $E(0,-\frac{\pi}{6})$:	pr(differ) $= 1/4$	The sequences will differ in about 1 in 4 places.
(iii) $E(\frac{\pi}{6},0)$:	pr(differ) $= 1/4$	The sequences will differ in about 1 in 4 places.
(iv) $E(\frac{\pi}{6},-\frac{\pi}{6})$:	pr(differ) $= 3/4$	The sequences will differ in about 3 in 4 places!

Recall that the sequence seen locally by A or B will, in every case, be uniformly random, in contrast to the angle-dependent *correlations* above.
(d)* Using (i) to (iv) above, argue (hmmm... mm... gosh!) that the local outcomes at A (resp. B) must be "instantaneously influenced" by the choice of angle at B (resp. A) i.e. that the correlations in the quantum measurement outcomes can only occur if there is some "spooky action at a distance" (Einstein's phrase) implied by the quantum rules for local operations on composite systems. Note also that by (a), although the "'instantaneous influence" must exist, it cannot be used to instantaneously send a signal from A to B (or vice versa) by suitable choice of local angle, since the effect is manifest only in *correlations* and not in any *local* measurement statistics.