PREREQUISITE MATERIAL

HT 2018

ARTUR EKERT

At a very instrumental level, quantum theory is a set of rules designed to answer questions such as 'given a specific preparation and a subsequent evolution compute probabilities for the outcomes of such and such test'. How do we represent preparations, evolutions and tests in mathematical terms, and how do we get probabilities? A mathematical setting for the quantum formalism is a vector space with an inner product. The result of any preparation is represented by some unit vector $|\psi\rangle$ and any test is represented by some other unit vector $|e\rangle$. The inner product of these two vectors, written as $\langle e|\psi\rangle$, gives the probability amplitude that an object prepared in state $|\psi\rangle$ will pass a test for being found in state $|e\rangle$. Probabilities are obtained by squaring absolute values of probability amplitudes, $|\langle e|\psi\rangle|^2$. After the test, in which the object was found to be in state $|e\rangle$, the object forgets about its previous state $|\psi\rangle$ and is indeed in state $|e\rangle$. This is the mysterious "quantum collapse", status of which we will discuss in the lectures. Mathematically we write this as $|e\rangle\langle e|\psi\rangle$ and read it from right to left. As we shall see in a moment $|e\rangle\langle e|$ describes a projection on state $|e\rangle$. If we let the initial state $|\psi\rangle$ evolve then right after the evolution U, which is represented by a unitary operator, the new state is $U|\psi\rangle$ and the probability amplitude that it will be found in state $|e\rangle$ is given by $\langle e|U|\psi\rangle$. The notation I have just used, full of $|\cdot\rangle$ and $\langle\cdot|$, follows the standard quantum notation that will likely be familiar to physicists, but may look odd to mathematicians or computer scientists. Love it or hate it, and I suggest the former, the 1939 Dirac bra-ket notation is so common that you simply have no choice but to learn it, especially if you want to study anything related to quantum theory. What follows is a basic introduction for the uninitiated.

$|e\rangle\langle e| \longleftarrow |\psi\rangle$

1. Vector spaces

1.1. I assume you are familiar with Euclidean vectors, those arrow-like geometric objects which are used to represent physical quantities such as velocities or forces. You know that any two velocities can be added to yield a third, and the multiplication of a velocity vector by a real number is another velocity vector. Thus a linear combination of vectors is another vector. Mathematicians simply took these properties and defined vectors as *anything* that can be added and multiplied by a number. This is basically what an Italian mathematician Giuseppe Peano (1858 - 1932) did in a chapter of his 1888 book with an impressive title, *Calcolo geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle operazioni della logica deduttiva*.

1.2. Vector spaces. Following Peano we define a vector space as a mathematical structure in which the notion of linear combination make sense. More formally, a complex vector space is a set V such that, given any two vectors $|a\rangle$ and $|b\rangle$ (that is, two elements of V) and any two complex numbers α and β, we can form the linear combination $\alpha|a\rangle + \beta|b\rangle$ which is also a vector in V. A subspace of V is any subset of V which is closed under vector addition and multiplication by complex numbers. Here I start using the Dirac bra-ket notation and write vectors in a somewhat fancy way as $|\text{label}\rangle$, where the "label" is anything that serves to specify what the vector is, for example, $|\uparrow\rangle$ and $|\downarrow\rangle$ may refer to an electron with spin up or down along some prescribed direction and $|0\rangle$ and $|1\rangle$ may describe a quantum bit, a qubit, holding either logical 0 or 1. These are often called 'ket' vectors or simply 'kets'. We will deal with 'bras' in a moment. A basis in V is a collection of vectors

Linear combinations must obey certain natural rules. Addition of vectors must be commutative and associative, with an identity (the zero vector, which will always be written as $\mathbf{0}$) and an inverse for each $|v\rangle$ (written as $-|v\rangle$). Multiplication by complex numbers must obey the two distributive laws: $(\alpha+\beta)|v\rangle = \alpha|v\rangle + \beta|v\rangle$ and $\alpha(|v\rangle + |w\rangle) = \alpha|v\rangle + \alpha|w\rangle$.

$|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle$ such that every vector $|v\rangle$ in V can be written in exactly one way as a linear combination of the basis vectors; $|v\rangle = \sum_i v_i |e_i\rangle$. The number of elements in a basis is called the dimension of V. The most common n-dimensional complex vector space is the space of ordered n-tuples of complex numbers, usually written as column vectors.

$$|a\rangle = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad |b\rangle = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \alpha|a\rangle + \beta|b\rangle = \begin{bmatrix} \alpha a_1 + \beta b_1 \\ \alpha a_2 + \beta b_2 \\ \vdots \\ \alpha a_n + \beta b_n \end{bmatrix} \tag{1}$$

In fact this is the space we will use most of the time. Throughout the course we will deal only with vector spaces of finite dimensions. This is sufficient for all our purposes and we will avoid many mathematical subtleties associated with infinite dimensional spaces.

1.3. Bras and kets. With any physical system we associate a complex vector space with an inner product, known as a Hilbert space $\mathcal{H}$. The inner product between vectors $|u\rangle$ and $|v\rangle$ in $\mathcal{H}$ is written as

$$\langle u | v \rangle$$

The inner product is a function that assigns to each pair of vectors $|u\rangle, |v\rangle \in \mathcal{H}$ a complex number $\langle u | v \rangle$ and satisfies the following conditions:

- $\langle u | v \rangle = \langle v | u \rangle^\star$,
- $\langle v | v \rangle \geq 0$ for all v,
- $\langle v | v \rangle = 0$ if and only if $v = 0$.

The inner product must be linear in the second argument but antilinear in the first argument;
$\langle c_1 u_1 + c_2 u_2 | v \rangle = c_1^\star \langle u_1 | v \rangle + c_2^\star \langle u_2 | v \rangle$ for any complex constants c_1 and c_2.

For example, for column vectors $|u\rangle$ and $|v\rangle$ in $\mathbb{C}^n$,

$$|u\rangle = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad |v\rangle = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{2}$$

the inner product is defined as

$$\langle u | v \rangle = u_1^\star v_1 + u_2^\star v_2 + \ldots + u_n^\star v_n. \tag{3}$$

Following Dirac we may split the inner product into two ingredients

$$\langle u | v \rangle \longrightarrow \langle u | \quad | v \rangle .$$

Here $|v\rangle$ is a ket vector and $\langle u|$ is called a bra vector, or a bra, and can be represented by a row vector

$$\langle u | = [u_1^\star,\ u_2^\star,\ \ldots,\ u_n^\star] .$$

The term "Hilbert space" used to be reserved for an infinite-dimensional inner product space that is complete i.e. every Cauchy sequence in the space converges to an element in the space. Nowadays, as in these notes, the term includes finite-dimensional spaces, which automatically satisfy the condition of completeness.

The inner product can now be viewed as the result of the matrix multiplication

$$\langle u | v \rangle = [u_1^\star,\ u_2^\star,\ \ldots,\ u_n^\star] \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1^\star v_1 + u_2^\star v_2 + \ldots + u_n^\star v_n.$$

Bras are vectors, for you can add them and multiply by scalars (here complex numbers), but they are vectors in the space $\mathcal{H}^\star$ which is dual to $\mathcal{H}$. Elements of $\mathcal{H}^\star$ are linear functionals, that is, linear maps from $\mathcal{H}$ to $\mathbb{C}$. Linear functional $\langle u |$ acting on any vector $|v\rangle$ in $\mathcal{H}$ gives a complex number $\langle u | v \rangle$.

All Hilbert spaces of the same dimension are isomorphic, so the differences between quantum systems cannot be really understood without additional structure. This structure is provided by a specific algebra of operators acting on $\mathcal{H}$.

1.4. Although $\mathcal{H}$ and $\mathcal{H}^\star$ are not identical spaces – the former is inhabited by kets and the latter by bras – they are closely related. There is a one-to-one map from one to the other, $|v\rangle \leftrightarrow \langle v|$, denoted by a dagger

"Is this a † which I see before me..."

$$\langle v | = (|v\rangle)^\dagger, \qquad |v\rangle = (\langle v|)^\dagger. \tag{4}$$

We usually omit the parentheses when it is obvious what the dagger operation applies to. The dagger operation, also known as Hermitian conjugation, is antilinear,

$$\begin{aligned} (c_1 |v_1\rangle + c_2 |v_2\rangle)^\dagger &= c_1^\star \langle v_1| + c_2^\star \langle v_2| , \\ (c_1 \langle v_1| + c_2 \langle v_2|)^\dagger &= c_1^\star |v_1\rangle + c_2^\star |v_2\rangle . \end{aligned}$$

When applied twice, the dagger operation is the identity map. In the matrix representation,

$$|\,v\rangle = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \overset{\dagger}{\longleftrightarrow} \langle v\,| = [v_1^\star,\ v_2^\star,\ \ldots,\ v_n^\star]\,.$$

Recall that the conjugate transpose or the Hermitian conjugate of an $n \times m$ matrix A is an $m \times n$ matrix $A^\dagger$, obtained by interchanging the rows and columns of A and taking complex conjugates of each entry in A, i.e. $A^\dagger_{ij} = A^\star_{ji}$. In mathematics texts it is often denoted by * rather than †.

1.5. The inner product brings geometry: the length or a norm of $|\,v\rangle$ is given by $||v|| = \sqrt{\langle\, v\,|\,v\,\rangle}$ and we call $|\,u\rangle$ and $|\,v\rangle$ orthogonal if $\langle\, u\,|\,v\,\rangle = 0$. Any maximal set of pairwise orthogonal vectors of unit length $\langle\, e_i\,|\,e_j\,\rangle = \delta_{ij}$ forms an orthonormal basis and any vector can be expressed as a linear combination of the basis vectors,

δ_{ij}, known as the "Kronecker delta", is a symbol that is defined to be zero for $i \neq j$ and to be one for $i = j$.

$$|\,v\rangle = \sum_i v_i\,|\,e_i\rangle\,, \quad \text{where} \quad v_i = \langle\, e_i\,|\,v\,\rangle\,.$$

Bras $\langle e_i\,|$ form the dual basis,

$$\langle v\,| = \sum_i v_i^\star\,\langle e_i\,|\,, \quad \text{where} \quad v_i^\star = \langle\, v\,|\,e_i\,\rangle\,.$$

To make the notation a bit less cumbersome we will sometimes label the basis kets as $|\,i\rangle$ rather than $|\,e_i\rangle$ and write

Do not confuse $|\,0\rangle$ with the zero vector. We will never write the zero vector as $|\,0\rangle$, it will be always written as 0 without any bra or ket decorations, e.g. $|\,v\rangle + 0 = |\,v\rangle$.

$$|\,v\rangle = \sum_i |\,i\rangle\,\langle\, i\,|\,v\,\rangle\,, \qquad \langle v\,| = \sum_j \langle\, v\,|\,i\,\rangle\,\langle i\,|\,.$$

With any isolated quantum system, which can be prepared in n perfectly distinguishable states, we can associate a Hilbert space $\mathcal{H}$ of dimension n such that each vector $|\,v\rangle \in \mathcal{H}$ of unit length, $\langle\, v\,|\,v\,\rangle = 1$, represents a quantum state of the system. The overall phase of the vector has no physical significance: $|\,v\rangle$ and $e^{i\varphi}\,|\,v\rangle$, for any real φ, describe the same state. The inner product $\langle\, u\,|\,v\,\rangle$ is the probability amplitude that a quantum system prepared in state $|\,v\rangle$ will be found in state $|\,u\rangle$. States corresponding to orthogonal vectors, $\langle\, u\,|\,v\,\rangle = 0$, are perfectly distinguishable for the system prepared in state $|\,v\rangle$ will never be found in state $|\,u\rangle$, and vice versa. In particular, states forming orthonormal bases are always perfectly distinguishable from each other. Choosing such states, as we shall see in a moment, is equivalent to choosing a particular quantum measurement.

1.6. Exercises.

(1) Show that both $\frac{1}{\sqrt{2}}\,(|\,0\rangle \pm |\,1\rangle)$ and $\frac{1}{\sqrt{2}}\,(|\,0\rangle \pm i\,|\,1\rangle)$ form orthonormal bases of a qubit. The qubit is prepared in state $\frac{1}{\sqrt{2}}\,\big(|\,0\rangle + e^{i\varphi}\,|\,1\rangle\big)$. What is the probability amplitude that right after the preparation the qubit will be found in state $\frac{1}{\sqrt{2}}\,(|\,0\rangle + |\,1\rangle)$? What is the corresponding probability?

(2) A vector space which is complete with respect to a norm $||\cdot||$, i.e. if every Cauchy sequence converges, is called a Banach space. A Hilbert space is a Banach space with a norm which is determined by an inner product. While a Hilbert space is always a Banach space, the converse does not hold; there are norms which are not given by an inner product. For example, the p-norm:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}},$$

where $p \geq 1$ and x_i are the components of vector $|\, x\rangle$, is the inner product norm iff $p = 2$. Show that any inner product norm satisfies the parallelogram law,

$$2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2,$$

for any two vectors $|\, x\rangle$ and $|\, y\rangle$. It turns out that given an arbitrary norm, there exists an inner product that induces that norm if and only if the norm satisfies the parallelogram law. Show that in this case

$$\langle\, x \,|\, y \,\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 - i\|x+iy\|^2 + i\|x-iy\|^2\right).$$

In plane geometry the parallelogram law simply states that the sum of squares of the diagonals of a parallelogram equals the sum of squares of its four sides.

Remember that the inner product is antilinear (or conjugate-linear) in the first argument, e.g. the inner product of $i\,|\, x\rangle$ and $|\, y\rangle$ is $-i\,\langle\, x \,|\, y \,\rangle$.

(3) Show that for any $|\, x\rangle, |\, y\rangle \in \mathcal{H}$

$$|\,\langle\, x \,|\, y \,\rangle\,| \leq \|x\| \cdot \|y\|$$

with equality if and only if $|\, x\rangle$ and $|\, y\rangle$ are linearly dependent. This is the Cauchy–Schwarz inequality, one of the most important inequalities in mathematics (no, I am not exaggerating).

(4) A polarising filter P_u transmits photons in state $|\, v\rangle$ with probability amplitude $\langle\, u \,|\, v \,\rangle$. Suppose you are told that each incoming photon is prepared in one of the two states, either $|\, a\rangle$ or $|\, b\rangle$, with equal probability. You can place a photodetector behind the filter and you can rotate the filter to choose $|\, u\rangle$. If $\langle\, a \,|\, b \,\rangle = 0$ you can reliably distinguish between the two preparations. Describe how this can be done and why the method fails when $\langle\, a \,|\, b \,\rangle \neq 0$.

2. Operators

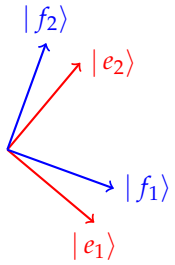
A linear map between two vector spaces $\mathcal{H}$ and $\mathcal{K}$ is a function $A : \mathcal{H} \to \mathcal{K}$ which respects linear combinations $A(c_1 \,|\, v_1\rangle + c_2 \,|\, v_2\rangle) = c_1 A \,|\, v_1\rangle + c_2 A \,|\, v_2\rangle$, for any vectors $|\, v_1\rangle, |\, v_2\rangle$ and any complex numbers c_1, c_2. We will focus mostly on endomorphisms, that is, on maps of $\mathcal{H}$ into itself, and we will call them operators. The symbol $\mathbb{1}$ is reserved for the identity operator that maps every element of $\mathcal{H}$ to itself. The product AB of two operators A and B is the operator obtained by first applying B to some ket $|\, v\rangle$ and then A to the ket which results from applying B: $A(B \,|\, v\rangle) = AB \,|\, u\rangle$. The order does matter for in general $AB \neq BA$. In the exceptional case in which $AB = BA$ one says that these two operators commute. The inverse of A, written as A^{-1}, is the operator which satisfies $AA^{-1} = \mathbb{1} = A^{-1}A$. For finite-dimensional spaces one only needs to check one of the two conditions, for any of the two implies the other, whereas on an infinite-dimensional space both must be checked. Finally, given a particular basis, operator A is uniquely determined by its matrix elements defined as $A_{ij} = \langle i \,|\, A \,|\, j\rangle$. The adjoint, or Hermitian conjugate, of A, denoted by $A^\dagger$, is defined by the relation

$$\langle i \,|\, A^\dagger \,|\, j\rangle = \langle j \,|\, A \,|\, i\rangle^\star, \quad \text{for all } |\, i\rangle, |\, j\rangle \in \mathcal{H}. \tag{5}$$

Operators for which $A^\dagger = A$, are called Hermitian or self-adjoint and operators for which $A^\dagger = A^{-1}$ are called unitary.

$$|\, a\rangle^\dagger = \langle a \,|\,, \ \langle a \,|^\dagger = |\, a\rangle$$
$$(\alpha \,|\, a\rangle + \beta \,|\, b\rangle)^\dagger = \alpha^\star \langle a \,|\, + \beta^\star \langle b \,|\,$$
$$(|\, a\rangle \langle b \,|\,)^\dagger = |\, b\rangle \langle a \,|\,$$
$$(AB)^\dagger = B^\dagger A^\dagger$$
$$(\alpha A + \beta B)^\dagger = \alpha^\star A^\dagger + \beta^\star B^\dagger$$
$$(A^\dagger)^\dagger = A$$

2.1. Outer products or dyads. Apart from the inner product $\langle u \,|\, v\rangle$, which is a complex number, we can also form the outer product $|\, u\rangle \langle v \,|$ which is a linear map (operator) on $\mathcal{H}$ or on $\mathcal{H}^\star$, depending how you look at it. This is what physicists like and what mathematicians hate about Dirac notation – a certain degree of healthy ambiguity. The result of $|\, u\rangle \langle v \,|$ acting on ket $|\, x\rangle$ is $|\, u\rangle \langle v \,|\, x\rangle$, that is vector $|\, u\rangle$ multiplied by the complex number $\langle v \,|\, x\rangle$. By the same token, the result of $|\, u\rangle \langle v \,|$ acting on bra $\langle y \,|$ is $\langle y \,|\, u\rangle \langle v \,|$, that is functional $\langle v \,|$ multiplied by the complex number $\langle y \,|\, u\rangle$. The product of two maps, $A = |\, a\rangle \langle b \,|$ followed by $B = |\, c\rangle \langle d \,|$, is a linear map BA, which in Dirac notation can be written as $BA = |\, c\rangle \langle d \,|\, a\rangle \langle b \,| = \langle d \,|\, a\rangle \,|\, c\rangle \langle b \,|$, that is, inner product (complex number) $\langle d \,|\, a\rangle$ times outer product (liner map) $|\, c\rangle \langle b \,|$.



$A = |\, f_1\rangle \langle e_1 \,| + |\, f_2\rangle \langle e_2 \,|$

Any operator on $\mathcal{H}$ can be expressed as a sum of outer products. Given an orthonormal basis $\{|\, e_i\rangle\}$, any operator which maps the basis vectors $|\, e_i\rangle$ into vectors $|\, f_i\rangle$ can be written as $\sum_i |\, f_i\rangle \langle e_i \,|$, where the sum is over all the vectors in the orthonormal basis. If vectors $\{|\, f_i\rangle\}$ also form an orthonormal basis then the operator simply "rotates" one orthonormal basis into another. These are unitary operators which preserve the inner product. In particular, if each $|\, e_i\rangle$ is mapped into $|\, e_i\rangle$, we obtain the identity operator

$$\sum_i |\, e_i\rangle \langle e_i \,| = \mathbb{1}.$$

This relation holds for *any* orthonormal basis and it is one of the most ubiquitous and useful formulas in quantum theory. For example, for any vector $|\, v\rangle$ and for any orthonormal basis $\{|\, e_i\rangle\}$ we have

$$|\, v\rangle = \mathbb{1} \,|\, v\rangle = \sum_i |\, e_i\rangle \langle e_i \,|\, v\rangle = \sum_i v_i \,|\, e_i\rangle, \tag{6}$$

where $v_i = \langle e_i \,|\, v\rangle$ are the components of $|\, v\rangle$. Finally, note that the adjoint of $|\, a\rangle \langle b \,|$ is $|\, b\rangle \langle a \,|$.

2.2. Trace. The trace is an operation which turns outer products into inner products,

$$|\, b\rangle \langle a \,| \longrightarrow \langle a \,|\, b\rangle. \tag{7}$$

We have just seen that any linear operator can be written as a sum of outer products, hence we can extend the definition of trace (by linearity) to any operator. Alternatively, for any square matrix A the trace of A is defined to be the sum

of its diagonal elements, $\mathrm{Tr}\, A = \sum_k \langle e_k \,|\, A \,|\, e_k \rangle = \sum_k A_{kk}$. You can show, using this definition or otherwise, that the trace is cyclic, $\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$ and linear $\mathrm{Tr}\,(\alpha A + \beta B) = \alpha \mathrm{Tr}\, A + \beta \mathrm{Tr}\, B$, where A and B are square matrices and α and β complex numbers. Moreover,

$\mathrm{Tr}\,(\alpha A + \beta B) = \alpha \mathrm{Tr}\, A + \beta \mathrm{Tr}\, B$
$\mathrm{Tr}\, |\, a\rangle \langle b\, | = \langle\, b \,|\, a\,\rangle$
$\mathrm{Tr}\, ABC = \mathrm{Tr}\, CAB = \mathrm{Tr}\, BCA$

$$\mathrm{Tr}\, |\, b\rangle \langle a\, | = \sum_k \langle\, e_k \,|\, b\,\rangle \langle\, a \,|\, e_k\,\rangle = \sum_k \langle\, a \,|\, e_k\,\rangle \langle\, e_k \,|\, b\,\rangle = \langle\, a \,|\, \mathbb{1}\,\rangle \,|\, b\rangle = \langle\, a \,|\, b\,\rangle\,. \quad (8)$$

Here the second term can be viewed both as the sum of the diagonal elements of $|\, b\rangle \langle a\,|$ in the $|\, e_k\rangle$ basis and as the sum of the products of two complex numbers $\langle\, e_k \,|\, b\,\rangle$ and $\langle\, a \,|\, e_k\,\rangle$. We have used the decomposition of the identity, $\sum_k |\, e_k\rangle \langle e_k\,| = \mathbb{1}$. Given that we can decompose the identity by choosing any orthonormal basis it is clear that the trace does not depend on the choice of the basis.

2.3. Examples. In this course we will often work with a 2-dimensional Hilbert space with a chosen orthonormal basis denoted as $|\, 0\rangle, |\, 1\rangle$, called the computational basis. Given

$$|\, 0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |\, 1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad |\, a\rangle = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \quad |\, b\rangle = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \quad (9)$$

we can write $|\, a\rangle = a_0\,|\, 0\rangle + a_1\,|\, 1\rangle$ and $|\, b\rangle = b_0\,|\, 0\rangle + b_1\,|\, 1\rangle$. We use the conjugate transpose (the dagger operation) to obtain the corresponding bras,

$$\langle 0\,| = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \langle 1\,| = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad \langle a\,| = \begin{bmatrix} a_0^\star & a_1^\star \end{bmatrix} \quad \langle b\,| = \begin{bmatrix} b_0^\star & b_1^\star \end{bmatrix} \quad (10)$$

and write the inner product of $|\, a\rangle$ and $|\, b\rangle$ explicitly as

$$\langle\, a \,|\, b\,\rangle = a_0^\star b_0 + a_1^\star b_1. \quad (11)$$

The standard matrix multiplication gives

$$|\, a\rangle \langle b\,| = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \begin{bmatrix} b_0^\star & b_1^\star \end{bmatrix} = \begin{bmatrix} a_0 b_0^\star & a_0 b_1^\star \\ a_1 b_0^\star & a_1 b_1^\star \end{bmatrix}. \quad (12)$$

The trace of $|\, a\rangle \langle b\,|$ is the sum of diagonal elements $a_0 b_0^\star + a_1 b_1^\star$, which is indeed equal to $\langle\, a \,|\, b\,\rangle$. Outer products which involve only vectors from the computational basis can be explicitly written as

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

As you can see, any 2×2 matrix can be expressed as a linear combination of the four matrices above,

$$\begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} = A_{00}\,|\, 0\rangle \langle 0\,| + A_{01}\,|\, 0\rangle \langle 1\,| + A_{10}\,|\, 1\rangle \langle 0\,| + A_{11}\,|\, 1\rangle \langle 1\,|\,.$$

In general, outer products which involve only vectors from an orthonormal basis, $\{|\, i\rangle\}$, are often used to express operators in that basis; operator A can be written as $A = \sum_{ij} A_{ij}\,|\, i\rangle \langle j\,|$. In the matrix language $|\, i\rangle \langle j\,|$ represents a matrix with all entries equal to zero except entry (i, j) which is equal to one. In particular, $|\, i\rangle \langle i\,|$ add up to the identity operator $\sum_i |\, i\rangle \langle i\,| = \mathbb{1}$, as we have already noticed.

2.4. Quantum evolutions. Any physically admissible evolution of an isolated quantum system is represented by a unitary operator. Please note that unitary operators preserve the inner product. Given unitary operator U and $|\, a'\rangle = U\,|\, a\rangle$, $|\, b'\rangle = U\,|\, b\rangle$ we have $\langle a'\,| = \langle a\,|\, U^\dagger$ and

$$\langle\, a' \,|\, b'\,\rangle = \langle a\,|\, U^\dagger U\,|\, b\rangle = \langle a\,|\, \mathbb{1}\,|\, b\rangle = \langle\, a \,|\, b\,\rangle\,. \quad (13)$$

Preserving the inner product implies preserving the norm induced by this product, that is, unit state vectors are mapped into unit state vectors, i.e. unitary operations are the isometries of the Euclidean norm.

2.5. Schrödinger equation. Unitary operators describing evolutions of quantum systems are usually derived from the Schrödinger equation,

$$\frac{d}{dt}\,|\,\psi(t)\rangle = -\frac{i}{\hbar}\,H\,|\,\psi(t)\rangle\,, \tag{14}$$

Here $\hbar = 1.05 \times 10^{-34}$ J s denotes Planck's constant. Theorists will always choose to work with a system of units where $\hbar = 1$.

where H is a Hermitian operator called the Hamiltonian. It contains a complete specification of all interactions both within the system and between the system and the external potentials. For time independent Hamiltonians the formal solution of the Schrödinger equation reads

$$|\,\psi(t)\rangle = U(t)\,|\,\psi(0)\rangle \qquad \text{where} \qquad U(t) = e^{-\frac{i}{\hbar}Ht} \tag{15}$$

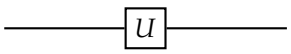
Any unitary matrix can be represented as the exponential of some Hermitian matrix, H and a real coefficient t,

We shall ignore the convergence issues

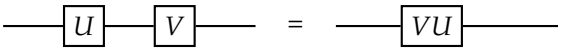
$$e^{itH} \equiv \mathbb{1} + itH + \frac{(it)^2}{2}H^2 + \frac{(it)^3}{2\cdot 3}H^3 ... = \sum_{n=0}^{\infty}\frac{(it)^n}{n!}\,H^n. \tag{16}$$

The state vector changes smoothly; for $t = 0$ the time evolution operator is merely the unit operator $\mathbb{1}$, and when t is very small $U(t) \approx \mathbb{1} - itH$ is close to the unit operator, differing from it by something of order t.

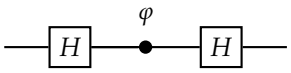
2.6. Quantum circuits. In this course we will hardly refer to the Schrödinger equation, instead we will assume that our clever colleagues, experimental physicists, are able to implement certain unitary operations and we will use these unitaries, like lego blocks, to construct other, more complex, unitaries. We refer to preselected elementary quantum operations as quantum logic gates and we often draw diagrams, called quantum circuits, to illustrate how they act on qubits. For example, unitary U acting on a single qubit is represented as



This diagram should be read from left to right. The horizontal line represents a qubit that is inertly carried from one quantum operation to another. A circuit composed of two gates, say U followed by V, is equivalent to a circuit composed of one gate described by the matrix product VU (note the order in which we multiply the matrices),



Now, here comes the most important sequence of single qubit gates: the Hadamard gate, followed by a phase shift gate, and followed by the Hadamard gate,



$$H = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Hadamard gate

This circuit represents a single qubit interference. You will see it over and over again, for it is quantum interference that gives quantum computation additional capabilities. The Hadamard gate H, in the Dirac notation, is written as

$$H = \frac{1}{\sqrt{2}}\left[\,|\,0\rangle\,\langle 0\,| + |\,0\rangle\,\langle 1\,| + |\,1\rangle\,\langle 0\,| - |\,1\rangle\,\langle 1\,|\right], \tag{17}$$

$$P_\varphi = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{bmatrix}$$

phase gate

and the phase gate, P_φ, as

$$P_\varphi = \left[\,|\,0\rangle\,\langle 0\,| + e^{i\varphi}\,|\,1\rangle\,\langle 1\,|\right]. \tag{18}$$

The qubit is usually prepared in state $|\,0\rangle$ and we are interested in probabilities of finding it in one of the basis states, $|\,0\rangle$ or $|\,1\rangle$, at the output.

The amplitude that input $|0\rangle$ will evolve into $|x\rangle$ is $\langle x|HP_\varphi H|0\rangle$, where $x = 0, 1$. This can be written as

$$\langle x|HP_\varphi H|0\rangle = \langle x|H\mathbb{1}P_\varphi \mathbb{1}H|0\rangle = \sum_{i,j} \langle x|H|j\rangle \langle j|P_\varphi|i\rangle \langle i|H|0\rangle .$$

After inserting the required entries from H and P_φ we obtain

$$\langle 0|HP_\varphi H|0\rangle = \cos\frac{\varphi}{2}, \quad \langle 1|HP_\varphi H|0\rangle|0\rangle = -i\sin\frac{\varphi}{2},$$

which means

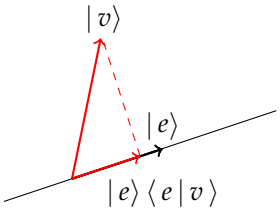
$$|0\rangle \to \cos\frac{\varphi}{2}|0\rangle - i\sin\frac{\varphi}{2}|1\rangle .$$

2.7. Projectors. A projector is any operator P such that $P^2 = P$. It does not matter how many times you apply P, it will have the same result as applying it just once. This makes sense from a geometric viewpoint, projecting the projection gives you the same projection back again. Here we will deal only with Hermitian projection operators, $P = P^\dagger$, called orthogonal projectors but we shall call them simply projectors.

For any normalised vector $|e\rangle$ ($\langle e|e\rangle = 1$) the outer product $P_e = |e\rangle\langle e|$ is an orthogonal projector for it is self-adjoint $P_e^\dagger = P_e$ and satisfies $P_eP_e = P_e$. The latter can be seen very neatly when expressed in Dirac notation, $|e\rangle\langle e|e\rangle\langle e| = |e\rangle\langle e|$. The projector $|e\rangle\langle e|$ projects on the one-dimensional subspace spanned by $|e\rangle$, which is self-evident in Dirac notation: $(|e\rangle\langle e|)|v\rangle = |e\rangle\langle e|v\rangle$, where $\langle e|v\rangle$ is the component of $|v\rangle$ along $|e\rangle$. Thus, given the orthonormal basis $\{|e_i\rangle\}$, the expression

$$|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2| + ... |e_k\rangle\langle e_k|$$

is the projector onto the subspace spanned by $\{|e_1\rangle, |e_2\rangle, ..., |e_k\rangle\}$. If we include all the basis vectors we end up projecting onto the entire Hilbert space $\mathcal{H}$. The projector, that projects onto the entire Hilbert space is, of course, the identity operator $\mathbb{1}$, hence, let us write this yet again, $\sum_i |e_i\rangle\langle e_i| = \mathbb{1}$.



A complete measurement in quantum theory is determined by the choice of an orthonormal basis $\{|e_i\rangle\}$ in $\mathcal{H}$, and every such basis in principle represents a possible measurement. Given a quantum system in state $|v\rangle$, such that

$$|v\rangle = \sum_i |e_i\rangle\langle e_i|v\rangle$$

the measurement in the basis $\{|e_i\rangle\}$ gives outcome labelled by e_k with probability $|\langle e_k|v\rangle|^2$ and leaves the system in state $|e_k\rangle$. This is consistent with our interpretation of the inner product $\langle e_k|v\rangle$ as the probability amplitude that a quantum system prepared in state $|v\rangle$ will be found in state $|e_k\rangle$. State vectors forming orthonormal bases are perfectly distinguishable from each other, $\langle e_i|e_j\rangle = \delta_{ij}$, hence there is no ambiguity about the outcome.

2.8. Orthogonal subspaces. Vectors from any orthonormal basis satisfy the following two conditions

$$\langle i|j\rangle = \delta_{ij}.$$

The orthonormality condition. The basis consists of unit vectors which are pairwise orthogonal (δ_{ij} is the Kronecker delta).

$$\sum_i |i\rangle\langle i| = \mathbb{1}.$$

The completeness condition means that *any* vector in $\mathcal{H}$ can be expressed as the sum of orthogonal projections on $|i\rangle$.

The notion of orthogonality and completeness can be naturally extended to subspaces of $\mathcal{H}$. Subspaces $\mathcal{E}_1$ and $\mathcal{E}_2$ are orthogonal if $\langle v_1|v_2\rangle = 0$ for every $|v_1\rangle \in \mathcal{E}_1$

and $|\,v_2\rangle \in \mathcal{E}_2$. We say that $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ form the orthogonal decomposition of $\mathcal{H}$, written as,

$$\mathcal{H} = \mathcal{E}_1 \oplus \mathcal{E}_2 \oplus \ldots \oplus \mathcal{E}_n, \tag{19}$$

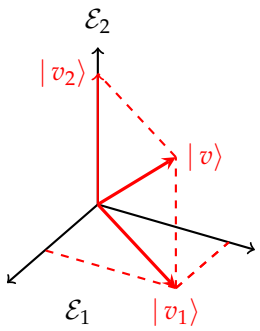
if the subspaces $\mathcal{E}_k$ are mutually orthogonal and any vector $|\,v\rangle \in \mathcal{H}$ has a unique representation

$$|\,v\rangle = |\,v_1\rangle + |\,v_2\rangle + \ldots + |\,v_n\rangle\,,$$

where $|\,v_k\rangle \in \mathcal{E}_k$. In geometric terms, vectors $|\,v_k\rangle$ are the results of orthogonal projections of vector $|\,v\rangle$ on the orthogonal subspaces $\mathcal{E}_k$: $P_k\,|\,v\rangle = |\,v_k\rangle$. The orthogonal decomposition of $\mathcal{H}$ can also be written in terms of projectors, as the decomposition of the identity into the sum of mutually orthogonal projectors:

$$\mathbb{1} = P_1 + P_2 + \ldots + P_n, \tag{20}$$

where projector P_k projects on subspace $\mathcal{E}_k$. This is a generalisation of the completeness relation, $\sum_i |\,i\rangle\,\langle i\,| = \mathbb{1}$, which used only projectors on one-dimensional subspaces spanned by vectors from orthonormal bases. Here, we have a collection of mutually orthogonal projectors $P_k P_l = P_k \delta_{kl}$, which form the decomposition of the identity $\sum_k P_k = \mathbb{1}$.



$$P_k P_l = P_k \delta_{kl}$$

Orthogonality condition for projectors.

$$\sum_k P_k = \mathbb{1}$$

Decomposition of the identity

They project on mutually orthogonal subspaces of $\mathcal{H}$ and any vector in $\mathcal{H}$ can be uniquely expressed as the sum of this orthogonal projections $|\,v\rangle = \sum_k P_k\,|\,v\rangle$.

So far we have identified measurements with orthonormal bases, or, if you wish, with a set of orthonormal projectors on the basis vectors. This is a complete measurement, which represents the best we can do in terms of resolving state vectors in the basis states. In general, for any decomposition of the identity $\sum_k P_k = \mathbb{1}$ into orthogonal projectors P_k ($P_k P_l = P_k \delta_{kl}$) there exists a measurement that takes a quantum system in state $|\,\psi\rangle$, outputs label k, with probability $\langle\psi\,|\,P_k\,|\,\psi\rangle$ and leaves the system in the state $P_k\,|\,\psi\rangle$ (multiplied by the normalisation factor i.e. divided by the length of $P_k\,|\,\psi\rangle$),

$$|\,\psi\rangle \longrightarrow \frac{P_k\,|\,\psi\rangle}{\sqrt{\langle\psi\,|\,P_k\,|\,\psi\rangle}}.$$

As an example of a measurement which is not complete consider a three dimensional Hilbert space and the following two orthogonal projectors $P = |\,1\rangle\,\langle 1\,| + |\,2\rangle\,\langle 2\,|$ and $Q = |\,3\rangle\,\langle 3\,|$ that form the decomposition of the identity. Suppose that a physical system is prepared in state $|\,v\rangle = c_1\,|\,1\rangle + c_2\,|\,2\rangle + c_3\,|\,3\rangle$. Ideally we would like to perform a complete measurement that would resolve the state $|\,v\rangle$ into the three basis states but suppose our experimental apparatus is not good enough and lumps together $|\,1\rangle$ and $|\,2\rangle$. It can only differentiate between the two subspaces associated with projectors P and Q. The apparatus, in this incomplete measurement, may find the system in the subspace associated with P. This happens with the probability $\langle v\,|\,P\,|\,v\rangle$, which is $|c_1|^2 + |c_2|^2$, and the state right after the measurement is the normalised $P\,|\,v\rangle$, that is,

$$\frac{c_1\,|\,1\rangle + c_2\,|\,2\rangle}{\sqrt{|c_1|^2 + |c_2|^2}}.$$

The measurement may also find the system in the subspace associated with Q with the probability $\langle v\,|\,Q\,|\,v\rangle$, which is $|c_3|^2$, resulting in the post-measurement state $|\,3\rangle$.

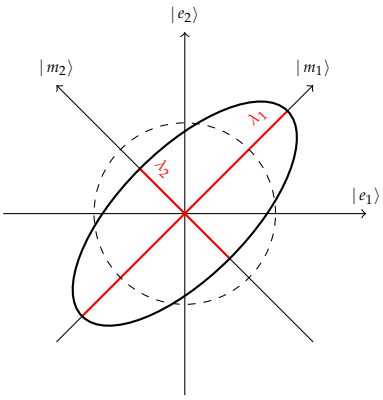
2.9. Spectral decomposition. An operator A is said to be normal if $AA^\dagger = A^\dagger A$. Both unitary and Hermitian operators are normal and all normal operators can be diagonalised by unitary matrices U. More precisely, M is normal if and only if there exists a unitary U such that

$$M = UDU^\dagger, \tag{21}$$

where D is the diagonal matrix, $D = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3 ...)$. The diagonal elements λ_j are known as the eigenvalues or the spectrum of M and the column vectors of U, which we can write as $|\, m_j\rangle = \sum_i U_{ij} |\, e_i\rangle$, are the corresponding eigenvectors of M, i.e. $M |\, m_j\rangle = \lambda_j |\, m_j\rangle$ and $\langle\, m_i |\, m_j\,\rangle = \delta_{ij}$, $\sum_j |\, m_j\rangle \langle m_j| = \mathbb{1}$. Thus any normal operator admits the spectral decomposition, $M = \sum_j \lambda_j |\, m_j\rangle \langle m_j|$. Some eigenvectors may share the same eigenvalue, which leads to a more general spectral decomposition,

$$M = \sum_k \lambda_k P_k,$$

where P_k projects on the subspace spanned by vectors that share eigenvalue λ_k.



Eigenvalues of Hermitian operators are real whereas for all unitary operators they are complex numbers of unit length: $\lambda_j = e^{i\alpha_j}$ for some real α_j. You may visualise it as follows: for each normal operator there exist a special basis (composed of its eignevectors) such that the operator kind of squeezes or stretches the space along each basis vector (or reflects, in the case of negative eigenvalues). This geometric picture is accurate for Hermitian operators, which have real eigenvalues, but also gives us some intuition for unitary operators, which preserve the length.

2.10. Exercises.

(1) The exponent of matrix A is defined as

$$e^A = \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + ... = \sum_{n=0}^{\infty} \frac{(A)^n}{n!}$$

Show that if H is self-adjoint (Hermitian), that is $H = H^\dagger$, then $U = e^{iHt}$ is unitary for any real t. Many quantum evolutions are expressed in this way. This is because matrix H, known as the Hamiltonian, is related to energies, which are measurable physical quantities, and t stands for time.

(2) Show that for any real α and for any A such that $A^2 = \mathbb{1}$

$$e^{i\alpha A} = \cos\alpha\, \mathbb{1} + i \sin\alpha\, A. \tag{22}$$

(3) A qubit (spin one-half particle) initially in state $|\, 0\rangle$ (spin up) is placed in a uniform magnetic field. The interaction between the field and the qubit is described by the Hamiltonian

$$H = \omega \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$$

where ω is proportional to the strength of the field. What is the state of the qubit after time $t = \pi/4\omega$?

In Earth's magnetic field, which is about 0.5 gauss, the value of ω is of the order of 10^6 cycles per second.

(4) Pauli operators

(5) The set of complex $N \times N$ matrices form a Hilbert space with the inner product $\frac{1}{N} \mathrm{Tr}\, A^\dagger B$. This inner product is often called the *Hilbert-Schmidt product*. In particular, the identity and the three Pauli operators form an orthonormal basis, with respect to the Hilbert-Schmidt product, in the space of complex 2×2 matrices. Any 2×2 matrix A can be expanded as $A = \frac{1}{2} \sum_{k=0}^{3} a_k \sigma_k$, where $a_k = \mathrm{Tr}\, \sigma_k A$. For self adjoint matrices, $A = A^\dagger$, the coefficients a_k are real.

3. Tensor products

Last but not least, we have tensor products. In quantum theory we use tensor products to construct Hilbert spaces associated with composed systems. Let states of system $\mathcal{A}$ be described by vectors in n dimensional Hilbert space $\mathcal{H}_A$ and states of system $\mathcal{B}$ by vectors in m dimensional Hilbert space $\mathcal{H}_B$. The combined system of $\mathcal{A}$ and $\mathcal{B}$ is then described by vectors in the nm dimensional tensor product space $\mathcal{H}_A \otimes \mathcal{H}_B$. Given bases $\{|\,a_1\rangle, \ldots, |\,a_n\rangle\}$ in $\mathcal{H}_A$ and $\{|\,b_1\rangle, \ldots, |\,b_m\rangle\}$ in $\mathcal{H}_B$ we form the tensor product basis consisting of the ordered pairs $|\,a_i\rangle \otimes \big|\,b_j\big\rangle$, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The tensor product space $\mathcal{H}_A \otimes \mathcal{H}_B$ consists of all linear combination of such tensor product basis vectors,

$$|\,\psi\rangle = \sum_{ij} c_{ij}\,|\,a_i\rangle \otimes \big|\,b_j\big\rangle\,.$$

Given $|\,a\rangle = \sum_i \alpha_i\,|\,a_i\rangle \in \mathcal{H}_A$ and $|\,b\rangle = \sum_j \beta_j\,\big|\,b_j\big\rangle \in \mathcal{H}_B$ we can write $|\,a\rangle \otimes |\,b\rangle = \sum_{ij} \alpha_i\beta_j\,|\,a_i\rangle\,\big|\,b_j\big\rangle$. In terms of column vectors, for example,

$$|\,a\rangle \otimes |\,b\rangle = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \otimes \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_0 \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \\ \alpha_1 \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \alpha_0\beta_0 \\ \alpha_0\beta_1 \\ \alpha_1\beta_0 \\ \alpha_1\beta_1 \end{bmatrix}.$$

Note that each element of the first vector multiplies the entire second vector. This is often the easiest way to get the tensor products in practice.

The tensor product operation $\otimes$ is distributive:

$$\begin{aligned} |\,a\rangle \otimes (\beta_1\,|\,b_1\rangle + \beta_2\,|\,b_2\rangle) &= \beta_1\,|\,a\rangle \otimes |\,b_1\rangle + \beta_2\,|\,a\rangle \otimes |\,b_2\rangle\,, \\ (\alpha_1\,|\,a_1\rangle + \alpha_2\,|\,a_2\rangle) \otimes |\,b\rangle &= \alpha_1\,|\,a_1\rangle \otimes |\,b\rangle + \alpha_2\,|\,a_2\rangle \otimes |\,b\rangle\,. \end{aligned}$$

We often drop the $\otimes$ symbol and simplify the labelling of the tensor product vectors from the computational basis. For example, a state of a quantum register composed of four qubits holding binary string 1001 may be is written as $|\,1\rangle \otimes |\,0\rangle \otimes |\,0\rangle \otimes |\,1\rangle$ or $|\,1\rangle\,|\,0\rangle\,|\,0\rangle\,|\,1\rangle$, or simply as $|\,1001\rangle$.

The bra corresponding to the tensor product state $|\,a\rangle \otimes |\,b\rangle$ is written as

$$(|\,a\rangle \otimes |\,b\rangle)^\dagger = \langle a\,| \otimes \langle b\,|$$

where the order of the factors on either side of $\otimes$ does not change when the dagger operation is applied. When we write tensor products we usually identify which vectors correspond to which subsystems by the order in which the respective tensor factors appear. If subsystem A is in state $|\,i\rangle$ and subsystem B in state $|\,j\rangle$ we write $|\,i\rangle \otimes |\,j\rangle$, or $|\,i\rangle\,|\,j\rangle$ (omitting the $\otimes$), or even $|\,ij\rangle$.

The tensor product of Hilbert spaces is a Hilbert space. The inner products on $\mathcal{H}_A$ and $\mathcal{H}_B$ give a natural inner product on $\mathcal{H}_A \otimes \mathcal{H}_B$. It is defined for any two product vectors $|\,a\rangle \otimes |\,b\rangle$ and $|\,a'\rangle \otimes |\,b'\rangle$ as

$$(\langle a\,| \otimes \langle b\,|)\,\big(\big|\,a'\big\rangle \otimes \big|\,b'\big\rangle\big) = \big\langle\, a\,\big|\,a'\,\big\rangle\,\big\langle\, b\,\big|\,b'\,\big\rangle\,,$$

and then extended by linearity to any two vectors in $\mathcal{H}_A \otimes \mathcal{H}_B$. If the bases $\{|\,a_i\rangle\}$ and $\{\big|\,b_j\big\rangle\}$ are orthonormal then so is the tensor product basis $\{|\,a_i\rangle \otimes \big|\,b_j\big\rangle\}$.

3.1. Quantum registers. Quantum computers store binary strings in registers composed of qubits. Consider a two-qubit quantum register. The standard (computational) basis of the two qubits taken as a composed system is given by the four product vectors $|\,0\rangle \otimes |\,0\rangle\,, |\,0\rangle \otimes |\,1\rangle\,, |\,1\rangle \otimes |\,0\rangle$ and $|\,1\rangle \otimes |\,1\rangle$. Here the order determines the subsystem; we have the first qubit and the second qubit (from left to right). We often drop the $\otimes$ symbol and write the standard product basis as $|\,0\rangle\,|\,0\rangle\,, |\,0\rangle\,|\,1\rangle\,, |\,1\rangle\,|\,0\rangle$ and $|\,1\rangle\,|\,1\rangle$, or as $|\,00\rangle\,, |\,01\rangle\,, |\,10\rangle$ and $|\,11\rangle$. Check that the following four states of two qubits,

$$\frac{1}{\sqrt{2}}\,(|\,00\rangle \pm |\,11\rangle)\,,\ \frac{1}{\sqrt{2}}\,(|\,01\rangle \pm |\,10\rangle)\,,$$

known as the Bell states, are normalised and pairwise orthogonal. They can be chosen as an orthonormal basis in the four-dimensional tensor product space.

Let us add one extra qubit to the register. The newly form register of size three can store individual binary strings such as,

$$|0\rangle \otimes |1\rangle \otimes |1\rangle \equiv |011\rangle, \tag{23}$$
$$|1\rangle \otimes |1\rangle \otimes |1\rangle \equiv |111\rangle, \tag{24}$$

but it can also store the two of them simultaneously. For if we take the first qubit and instead of setting it to $|0\rangle$ or $|1\rangle$ we prepare a superposition $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ then we obtain

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |1\rangle \otimes |1\rangle \equiv \frac{1}{\sqrt{2}}(|011\rangle + |111\rangle). \tag{25}$$

In fact we can prepare this register in a superposition of all eight binary strings it can hold – it is enough to put each qubit into the superposition $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. This gives

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{26}$$

which can also be written as

$$\sum_{x\in\{0,1\}^3} |x\rangle = |000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle. \tag{27}$$

Here we have dropped the normalisation constant $2^{-3/2}$. We will often do it for the clarity of the exposition.

In general quantum states of an n qubit register live in the n-fold tensor product of the two dimensional spaces. This n-fold tensor power, written as $\otimes^n\mathcal{H} = \mathcal{H} \otimes \ldots \otimes \mathcal{H}$, is a space of dimension 2^n with basis $|a_1\rangle \otimes \ldots \otimes |a_n\rangle$ $(a_1, \ldots, a_n = 0, 1)$ labelled by the 2^n n-bit strings $a_1 \ldots a_n$. We often write $|a_1\rangle \otimes \ldots \otimes |a_n\rangle$ simply as $|a_1 \ldots a_n\rangle$. When we bring together two registers, comprising of n and m qubits respectively, we form a new $n+m$ qubit register, $(\otimes^n\mathcal{H}) \otimes (\otimes^m\mathcal{H}) = \otimes^{n+m}\mathcal{H}$, with the computational basis $|a_1 \ldots a_n\rangle \otimes |b_1 \ldots b_n\rangle = |a_1 \ldots a_n b_1 \ldots b_n\rangle$.

This is how we label tensor products of vectors from the computational basis:
$|0\rangle \otimes |1\rangle = |01\rangle$
$|01\rangle \otimes |1\rangle = |011\rangle$
$|011\rangle \otimes |101\rangle = |011101\rangle$
$|11110\rangle \otimes |100\rangle = |11110100\rangle$

3.2. Entanglement. Even though any vector $|\psi\rangle$ in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as a *linear combination* of tensor product basis vectors, only some vectors in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written directly as tensor products, $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. They are called product vectors. Most vectors in $\mathcal{H}_A \otimes \mathcal{H}_B$ do not admit such a decomposition. Vectors that are not product vectors are called entangled. For example, any state of two qubits can be written as a linear combination of vectors from the computational basis $|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$; some of these states are separable, e.g.

$$\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$

and some of them are not, e.g. the most popular entangled states of two qubits, known as the Bell states,

$$\frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \ \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle),$$

do not admit any tensor product decomposition.

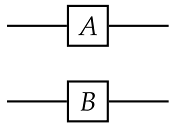
3.3. Operators revisited. We will also need the concept of the tensor product of two operators. If A is an operator on $\mathcal{H}_A$ and B an operator on $\mathcal{H}_B$ then the tensor product operator $A \otimes B$ is an operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ defined by its action on product vectors

$$(A \otimes B)\ (|a\rangle \otimes |b\rangle) = (A|a\rangle) \otimes (B|b\rangle).$$

Its action on all other vectors is determined by linearity,

$$A \otimes B \left(\sum_{ij} c_{ij} |a_i\rangle \otimes |b_j\rangle \right) = \sum_{ij} c_{ij}\ A|a_i\rangle \otimes B|b_j\rangle.$$

For example, two gates, A and B, acting in parallel on two different qubits,



are described by the tensor product $U \otimes V$.

The matrix elements of $A \otimes B$ are given by

$$\begin{aligned}(A \otimes B)_{ik,jl} &= \langle a_i \,|\, \langle b_k \,|\, (A \otimes B) \,\big|\, a_j \big\rangle \,|\, b_l \rangle = \langle a_i \,|\, A \,\big|\, a_j \big\rangle \langle b_k \,|\, B \,|\, b_l \rangle \\ &= A_{ij} B_{kl}.\end{aligned}$$

The composition of $(A' \otimes B')$ followed by $(A \otimes B)$ is written as the product

$$(A \otimes B)(A' \otimes B') = AA' \otimes BB'.$$

In particular, for outer products, we have

$$(|\, i\rangle \otimes |\, j\rangle)(\langle k \,| \otimes \langle l \,|) \equiv |\, i\rangle \langle k \,| \otimes |\, j\rangle \langle l \,|,$$

which we will often write as

$$|\, ij\rangle \langle kl \,| \equiv |\, i\rangle \langle k \,| \otimes |\, j\rangle \langle l \,|.$$

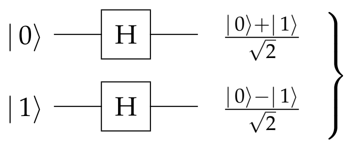
In practice we just form block diagonal matrices, e.g.,

$$A \otimes B = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \otimes \begin{bmatrix} B_{00} & B_{01} \\ B_{10} & B_{11} \end{bmatrix} = \left[\begin{array}{c|c} A_{00}B & A_{01}B \\ \hline A_{10}B & A_{11}B \end{array}\right],$$

where each element of the first matrix multiplies the entire second matrix,

$$\left[\begin{array}{c|c} A_{00}B & A_{01}B \\ \hline A_{10}B & A_{11}B \end{array}\right] = \left[\begin{array}{cc|cc} A_{00}B_{00} & A_{00}B_{01} & A_{01}B_{00} & A_{01}B_{01} \\ A_{00}B_{10} & A_{00}B_{11} & A_{01}B_{10} & A_{01}B_{11} \\ \hline A_{10}B_{00} & A_{10}B_{01} & A_{11}B_{00} & A_{11}B_{01} \\ A_{10}B_{10} & A_{10}B_{11} & A_{11}B_{10} & A_{11}B_{11} \end{array}\right].$$

The tensor product matrix has composite indices, $(A \otimes B)_{ik,jl}$, here $ik = 00, 01, 10, 11$ labels rows and $jl = 00, 01, 10, 11$ labels columns and we always use the lexicographical order, $00, 01, 10, 11$. For example, as you can see above, $(A \otimes B)_{01,11}$ is the entry in the second row and the fourth column and reads $A_{01}B_{11}$. For example, the circuit composed of two Hadamard gates acting in parallel ,



$$= \tfrac{1}{2}\left(|\, 00\rangle - |\, 01\rangle + |\, 10\rangle - |\, 11\rangle\right)$$

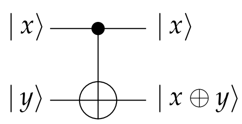
is described by the tensor product of the two Hadamard matrices, $H \otimes H$,

$$H \otimes H = \frac{1}{\sqrt{2}} \left[\begin{array}{c|c} H & H \\ \hline H & -H \end{array}\right] = \frac{1}{2} \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array}\right].$$

Can you see, by looking at this tensor product matrix, the result of $(H \otimes H)\,|\, 01\rangle$.

Most operators on $\mathcal{H}_A \otimes \mathcal{H}_B$ cannot be written directly as a tensor product of two operators on constituent subspaces. For example, take a controlled-NOT gate

$$\text{C-NOT} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right] \qquad (28)$$



where $x, y = 0$ or 1 and $\oplus$ denotes XOR or addition modulo 2. It is described by the matrix that does not admit any tensor product decomposition, but it can be written as the sum of tensor products

$$|\,0\rangle\langle 0\,| \otimes \mathbb{1} + |\,1\rangle\langle 1\,| \otimes (|\,0\rangle\langle 1\,| + |\,1\rangle\langle 0\,|).$$

3.4. Partial trace. If $A \otimes B$ is a tensor product operator on $\mathcal{H}_A \otimes \mathcal{H}_B$, the the partial trace over A or B is defined, respectively, as

$$\mathrm{Tr}_A\ A \otimes B = (\mathrm{Tr}\, A)B, \quad \mathrm{Tr}_B\ A \otimes B = A(\mathrm{Tr}\, B). \tag{29}$$

This definition is then extended to any operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ by linearity. For example, for any M on $\otimes^2\mathcal{H}$ (tensor product space associated with two qubits) with block form written in the standard basis $\{|\,00\rangle, |\,01\rangle, |\,10\rangle, |\,11\rangle\}$,

$$M = \left[\begin{array}{c|c} P & Q \\ \hline R & S \end{array}\right],$$

where P, Q, R, S are 2×2 sized sub-matrices, we have

$$\mathrm{Tr}_A\ M = P + Q, \quad \mathrm{Tr}_B\ M = \left[\begin{array}{c|c} \mathrm{Tr}\, P & \mathrm{Tr}\, Q \\ \hline \mathrm{Tr}\, R & \mathrm{Tr}\, S \end{array}\right].$$

The same holds for general M on any $\mathcal{H}_A \otimes \mathcal{H}_B$ with corresponding block form ($m \times m$ blocks of $n \times n$ sized sub-matrices, where m and n are the dimensions of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively).

3.5. Measuring subsystems. Consider a system AB consisting of two parts A and B, and suppose AB as a whole is in the state $|\,\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. Part A is now subject to a measurement defined by an orthonormal basis in $\mathcal{H}_A$, e.g. $\{|\,a_1\rangle, \ldots, |\,a_n\rangle\}$, where n is the dimension of $\mathcal{H}_A$. You can always write $|\,\psi\rangle$ as

$$|\,\psi\rangle = |\,a_1\rangle \otimes |\,v_1\rangle + \ldots + |\,a_n\rangle \otimes |\,v_n\rangle,$$

where $|\,v_1\rangle, \ldots, |\,v_n\rangle$ are unnormalised vectors in $\mathcal{H}_B$ (some of the $|\,v_i\rangle$'s may be zero vectors).

- The probability of the outcome $|\,a_i\rangle$ is $p_i = \langle\, v_i \,|\, v_i \,\rangle$.
- If the $|\,a_i\rangle$ outcome occurs, the final state of part A is $|\,a_i\rangle$ and the final state of part B is normalised $|\,v_i\rangle$, that is, $|\,v_i\rangle \,/\, \sqrt{\langle\, v_i \,|\, v_i \,\rangle}$.

This is equivalent to the following decomposition of the identity $\mathbb{1} \otimes \mathbb{1}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ into the orthogonal projectors,

$$(|\,a_1\rangle\langle a_1\,| + \ldots + |\,a_n\rangle\langle a_n\,|) \otimes \mathbb{1} = \sum_{i=1}^{n} |\,a_i\rangle\langle a_i\,| \otimes \mathbb{1} = \sum_{i=1}^{n} P_i,$$

where $P_i = |\,a_i\rangle\langle a_i\,| \otimes \mathbb{1}$. You can easily check that $(P_i)^2 = P_i$ and $P_i P_j = P_i \delta_{ij}$.

- The probability of the outcome $|\,a_i\rangle$ is $p_i = \langle\psi\,|\,P_i\,|\,\psi\rangle$.
- If the $|\,a_i\rangle$ outcome occurs, the final state of the system is

$$|\,\psi\rangle \longrightarrow \frac{P_i\,|\,\psi\rangle}{\sqrt{\langle\psi\,|\,P_i\,|\,\psi\rangle}} = |\,a_i\rangle \otimes \frac{|\,v_i\rangle}{\sqrt{\langle\, v_i \,|\, v_i \,\rangle}}$$

3.6. Example. As an example consider a pair of qubits in the state

$$|\,\psi\rangle = \frac{1}{\sqrt{2}}\,(|\,00\rangle + |\,11\rangle)\,.$$

We now perform the measurement on the first qubit in the basis

$$\{|\,a_1\rangle, |\,a_2\rangle\} = \left\{\frac{1}{\sqrt{2}}(|\,0\rangle + |\,1\rangle), \frac{1}{\sqrt{2}}(|\,0\rangle - |\,1\rangle)\right\}$$

The first method requires to express $|\,\psi\rangle$ as

$$|\,\psi\rangle = |\,a_1\rangle \otimes |\,v_1\rangle + |\,a_1\rangle \otimes |\,v_1\rangle\,.$$

You can easily show (I hope) that

$$|\,v_1\rangle = \frac{1}{\sqrt{2}}(|\,0\rangle + |\,1\rangle), \quad |\,v_2\rangle = \frac{1}{\sqrt{2}}(|\,0\rangle - |\,1\rangle).$$

Thus the probability of the outcome $|\,a_1\rangle$ is $\langle\, v_1 \,|\, v_1 \,\rangle = \frac{1}{2}^2 + \frac{1}{2}^2 = \frac{1}{2}$ and if the outcome $|\,a_1\rangle$ occurs the final state of the first qubit is $|\,a_1\rangle$ and the final state of the second qubit is $|\,v_1\rangle$ (which happens to be normalised in this particular case).

If you choose the second method then you write

$$P_1 = |\,a_1\rangle\,\langle a_1\,| \otimes \mathbb{1} = \frac{1}{2}\,(|\,0\rangle\,\langle 0\,| + |\,1\rangle\,\langle 1\,| + |\,0\rangle\,\langle 1\,| + |\,1\rangle\,\langle 0\,|) \otimes \mathbb{1},$$

evaluate

$$\begin{aligned} P_1\,|\,\psi\rangle &= \frac{1}{2\sqrt{2}}\,[(|\,0\rangle + |\,1\rangle) \otimes |\,0\rangle + (|\,0\rangle + |\,1\rangle) \otimes |\,1\rangle] \\ &= \frac{1}{2\sqrt{2}}\,(|\,0\rangle + |\,1\rangle) \otimes (|\,0\rangle + |\,1\rangle) \end{aligned}$$

and

$$\begin{aligned} \langle\psi\,|\,P_1\,|\,\psi\rangle &= \left[\frac{1}{\sqrt{2}}\,(\langle 0\,| \otimes \langle 0\,| + \langle 1\,| \otimes \langle 1\,|)\right] P_1 \left[\frac{1}{\sqrt{2}}\,(|\,0\rangle \otimes 0 + |\,1\rangle \otimes 1)\right] \\ &= \frac{1}{4}\,\langle\, 0\,|\,0\,\rangle + \frac{1}{4}\,\langle\, 1\,|\,1\,\rangle = \frac{1}{2}. \end{aligned}$$

Are you suspicious that something weird happened here? You should be. Assume the two qubits are light years apart. It seems that by making a measurement on the first qubit we have had an instantaneous effect on the second qubit. Outcomes $|\,a_1\rangle$ and $|\,a_2\rangle$ prepare the second qubit is states $|\,v_1\rangle$ and $|\,v_2\rangle$ respectively. Given that $\langle\, v_1 \,|\, v_2 \,\rangle = 0$ the two states can be reliably distinguished. Does it imply a possibility of instantaneous communication? No, it does not. Why not?

3.7. Exercises.

(1) Prove that the state $|\,\psi\rangle = a\,|\,00\rangle + b\,|\,01\rangle + c\,|\,10\rangle + d\,|\,11\rangle$ is entangled iff $ad - bc \neq 0$. Deduce that the state $\frac{1}{2}(|\,00\rangle + |\,01\rangle + |\,10\rangle + (-1)^k\,|\,11\rangle)$ is entangled for $k = 1$ and unentangled for $k = 0$. Express the latter case explicitly as a product state.