qitd114

Hilbert Space Quantum Mechanics

Robert B. Griffiths

Version of 16 January 2014

Contents

References:

CQT = *Consistent Quantum Theory* by Griffiths (Cambridge, 2002), Ch. 2; Ch. 3; Ch. 4 except for Sec. 4.3; Ch. 6.

QCQI = *Quantum Computation and Quantum Information* by Nielsen and Chuang (Cambridge, 2000). Secs. 1.1, 1.2; 2.1.1 through 2.1.7; 2.2.1

1 Introduction

1.1 Hilbert space

★ In quantum mechanics the state of a physical system is represented by a vector in a *Hilbert space*: a complex vector space with an inner product.

◦ The term "Hilbert space" is often reserved for an infinite-dimensional inner product space having the property that it is complete or closed. However, the term is often used nowadays, as in these notes, in a way that includes finite-dimensional spaces, which automatically satisfy the condition of completeness.

★ We will use *Dirac notation* in which the vectors in the space are denoted by $|v\rangle$, called a *ket*, where v is some symbol which identifies the vector.

One could equally well use something like $\vec{v}$ or $\mathbf{v}$. A multiple of a vector by a complex number c is written as $c|v\rangle$—think of it as analogous to $c\vec{v}$ of $c\mathbf{v}$.

★ In Dirac notation the inner product of the vectors $|v\rangle$ with $|w\rangle$ is written $\langle v|w\rangle$. This resembles the ordinary dot product $\vec{v}\cdot\vec{w}$ except that one takes a complex conjugate of the vector on the left, thus think of $\vec{v}^*\cdot\vec{w}$.

1.2 Qubit

★ The simplest interesting space of this sort is two-dimensional, which means that every vector in it can be written as a linear combination of two vectors which form a basis for the space. In quantum information the *standard* (or *computational*) basis vectors are denoted $|0\rangle$ and $|1\rangle$, and it is assumed that both of them are *normalized* and that they are mutually *orthogonal*

$$\langle 0|0\rangle = 1 = \langle w|w\rangle, \quad \langle 0|1\rangle = 0 = \langle 1|0\rangle. \tag{1}$$

(Note that $\langle v|w\rangle = \langle w|v\rangle^*$, so $\langle 0|1\rangle = 0$ suffices.)

• The notation $|0\rangle$ and $|1\rangle$ is intended to suggest an analogy, which turns out to be very useful, with an ordinary bit (binary digit) that takes the value 0 or 1. In quantum information such a two-dimensional Hilbert space, or the system it represents, is referred to as a *qubit* (pronounced "cubit"). However, there are disanalogies as well. Linear combinations like $0.3|0\rangle + 0.7i|1\rangle$ make perfectly good sense in the Hilbert space, and have a respectable physical interpretation, but there is nothing analogous for the two possible states 0 and 1 of an ordinary bit.

★ In quantum mechanics a two-dimensional complex Hilbert space $\mathcal{H}$ is used for describing the angular momentum or "spin" of a spin-half particle (electron, proton, neutron, silver atom), which then provides a physical representation of a qubit.

◦ The polarization of a photon (particle of light) is also described by $d = 2$, so represents a qubit.

• A state or vector $|v\rangle$ says something about one component of the spin of the spin half particle. The usual convention is:

$$|0\rangle = |z^+\rangle \leftrightarrow S_z = +1/2, \quad |1\rangle = |z^-\rangle \leftrightarrow S_z = -1/2, \tag{2}$$

where S_z, the z component of angular momentum is measured in units of $\hbar$. Here are some other correspondences:

$$\begin{aligned} \sqrt{2}|x^+\rangle = |0\rangle + |1\rangle \leftrightarrow S_x = +1/2, \quad \sqrt{2}|x^-\rangle = |0\rangle - |1\rangle \leftrightarrow S_x = -1/2, \\ \sqrt{2}|y^+\rangle = |0\rangle + i|1\rangle \leftrightarrow S_y = +1/2, \quad \sqrt{2}|y^-\rangle = |0\rangle - i|1\rangle \leftrightarrow S_y = -1/2. \end{aligned} \tag{3}$$

The general rule is that if w is a direction in space corresponding to the angles θ and ϕ in polar coordinates,

$$|0\rangle + e^{i\phi}\tan(\theta/2)|1\rangle \leftrightarrow S_w = +1/2, \quad |0\rangle - e^{i\phi}\cot(\theta/2)|1\rangle \leftrightarrow S_w = -1/2, \tag{4}$$

but see the comments below on normalization.

1.3 Intuitive picture

★ Physics consists of more than mathematics: along with mathematical symbols one always has a "physical picture," some sort of intuitive idea or geometrical construction which aids in thinking about what is going on in more approximate and informal terms than is possible using "bare" mathematics.

★ Most physicists think of a spin-half particle as something like a little top or gyroscope which is spinning about some axis with a well-defined direction in space, the direction of the angular momentum vector.

• This physical picture is often very helpful, but there are circumstances in which it can mislead, as can any attempt to visualize the quantum world in terms of our everyday experience. So one should be aware of its limitations.

• In particular, the axis of a gyroscope has a very precise direction in space, which is what makes such objects useful. But thinking of the spin of a spin-half particle as having a precise direction can mislead. A better (though by no means exact) physical picture is to think of the spin-half particle as having an angular momentum vector pointing in a random direction in space, but subject to the constraint that a particular component of the angular momentum, say S_z, is positive, rather than negative.

• Thus in the case of $|0\rangle$, which means $S_z = +1/2$, think of S_x and S_y as having random values. Strictly speaking these quantities are undefined, so one should not think about them at all. However, it is rather difficult to have a mental picture of an object spinning in three dimensions, but which has only one component of angular momentum. Thus treating one component as definite and the other two as random, while not an exact representation of quantum physics, is less likely to lead to incorrect conclusions than if one thinks of all three components as having well-defined values.

◦ An example of an incorrect conclusion is the notion that a spin-half particle can carry a large amount of information in terms of the orientation of its spin axis. To specify the orientation in space of the axis of a gyroscope requires on the order of $\log_2(1/\Delta\theta) + \log_2(1/\Delta\phi)$ bits, where $\Delta\theta$ and $\Delta\phi$ are the precisions with which the direction is specified (in polar coordinates). This can be quite a few bits, and in this sense the direction along which the angular momentum vector of a gyroscope is pointing can "contain" or "carry" a large amount of information. By contrast, the spin degree of freedom of a spin-half particle never carries or contains more than 1 bit of information, a fact which if ignored gives rise to various misunderstandings and paradoxes.

1.4 General d

★ A Hilbert space $\mathcal{H}$ of dimension $d = 3$ is referred to as a *qutrit*, one with $d = 4$ is sometimes called a *ququart*, and the generic term for any $d > 2$ is *qudit*. We will assume $d < \infty$ to avoid complications which arise in infinite-dimensional Hilbert spaces.

★ A collection of linearly independent vectors $\{|\beta_j\rangle\}$ form a *basis* of $\mathcal{H}$ provided any $|\psi\rangle$ in $\mathcal{H}$ can be written as a linear combination:

$$|\psi\rangle = \sum_j c_j |\beta_j\rangle. \tag{5}$$

The number d of vectors forming the basis is the *dimension* of $\mathcal{H}$ and does not depend on the choice of basis.

★ A particularly useful case is an *orthonormal basis* $\{|j\rangle\}$ for $j = 1, 2, \dots d$, with the property that

$$\langle j|k\rangle = \delta_{jk} : \tag{6}$$

the inner product of two basis vectors is 0 for $j \neq k$, i.e., they are *orthogonal*, and equal to 1 for $j = k$, i.e., they are *normalized*.

• If we write

$$|v\rangle = \sum_j v_j |j\rangle, \quad |w\rangle = \sum_j w_j |j\rangle, \tag{7}$$

where the coefficients v_j and w_j are given by

$$v_j = \langle j|v\rangle, \quad w_j = \langle j|w\rangle, \tag{8}$$

the inner product can be written as

$$(|v\rangle)^\dagger |w\rangle = \langle v|w\rangle = \sum_j v_j^* w_j, \tag{9}$$

which can be thought of as the product of a "bra" vector

$$\langle v| = (|v\rangle)^\dagger = \sum_j v_j^* \langle j| \tag{10}$$

with the "ket" vector $|w\rangle$. (The terminology goes back to Dirac, who referred to $\langle v|w\rangle$ as a *bracket*.)

• For more on the † operation, see below.

★ It is often convenient to think of $|w\rangle$ as represented by a column vector

$$|w\rangle = \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_d \end{pmatrix}, \tag{11}$$

and $\langle v|$ by a row vector

$$\langle v| = (v_1^*,\ v_2^*,\ \cdots v_d^*). \tag{12}$$

The inner product (9) is then the matrix product of the row times the column vector.

∘ Of course the numbers v_j and w_j depend on the basis. The inner product $\langle v|w\rangle$, however, is independent of the choice of basis.

1.5 Kets as physical properties

★ In quantum mechanics, two vectors $|\psi\rangle$ and $c|\psi\rangle$, where c is any *nonzero* complex number have *exactly the same* physical significance. For this reason it is sometimes helpful to say that the physical state corresponds not to a particular vector in the Hilbert space, but to the *ray*, or one-dimensional subspace, defined by the collection of all the complex multiples of a particular vector.

• One can always choose c (assuming $|\psi\rangle$ is not the zero vector, but that never represents any physical situation) in such a way that the $|\psi\rangle$ corresponding to a particular physical situation is *normalized*, $\langle\psi|\psi\rangle = 1$ or $\|\psi\| = 1$, where the *norm* $\|\psi\|$ of a state $|\psi\rangle$ is the positive square root of

$$\|\psi\|^2 = \langle\psi|\psi\rangle, \tag{13}$$

and is zero if and only if $|\psi\rangle$ is the (unique) zero vector, which will be written as 0 (and is not to be confused with $|0\rangle$).

∘ Normalized vectors can always be multiplied by a *phase factor*, a complex number of the form $e^{i\phi}$ where ϕ is real, without changing the normalization or the physical interpretation, so normalization by itself does not single out a single vector representing a particular physical state of affairs.

∘ For many purposes it is convenient to use normalized vectors, and for this reason some students of the subject have the mistaken impression that *any* vector representing a quantum system *must* be normalized. But that is to turn convenience into legalism. There are circumstances in which (as we shall see) it is more convenient not to use normalized vectors.

• The state of a single qubit is always a linear combination of the basis vectors $|0\rangle$ and $|1\rangle$,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{14}$$

where α and β are complex numbers. When $\alpha \neq 0$ this can be rewritten as

$$\alpha|0\rangle + \beta|1\rangle = \alpha\Big(|0\rangle + (\beta/\alpha)|1\rangle\Big) = \alpha\Big(|0\rangle + \gamma|1\rangle\Big), \quad \gamma := \beta/\alpha. \tag{15}$$

Since the physical significance of this state does not change if it is multiplied by a (nonzero) constant, we may multiply by α^{-1} and obtain a standard (unnormalized) form

$$|0\rangle + \gamma|1\rangle \tag{16}$$

characterized by a single complex number γ. There is then a one-to-one correspondence between different physical states or rays, and complex numbers γ, if one includes $\gamma = \infty$ to signify the ray generated by $|1\rangle$.

• Avoid the following mistake. Just because $|\psi\rangle$ and $c|\psi\rangle$ have the same physical interpretation does *not* mean that one can multiply a vector inside some formula by a constant without changing the physics. Thus $|\chi\rangle + |\psi\rangle$ and $|\chi\rangle - |\psi\rangle$ will (in general) *not* have the same physical significance. See the examples in (3) and the discussion above of (16). An overall constant makes no difference, but changing the relative magnitudes or phases of two kets in a sum can make a difference.

★ Two nonzero vectors $|\psi\rangle$ and $|\phi\rangle$ which are *orthogonal*, $\langle\phi|\psi\rangle = 0$, represent *distinct* physical properties: if one corresponds to a property, such as $S_z = +1/2$, which is a correct description of a physical system at a particular time, then the other corresponds to a physical property which is not true (false) for this system at this time. That is, the physical properties are *mutually exclusive*.

□ Exercise. There are six vectors in (2) and (3). Which pairs represent mutually exclusive properties?

∘ An example of mutually exclusive properties from classical physics: $P :=$ the energy is less than 1 Joule; $Q :=$ the energy is greater than 2 Joules.

★ There are cases in which $|\psi\rangle$ is neither a multiple of $|\phi\rangle$, nor is it orthogonal to $|\phi\rangle$. For example, the $S_z = +1/2$ vector in (2) and the $S_x = +1/2$ vector in (3). These represent neither the same physical situation, nor do they represent distinct physical situations. Instead they represent *incompatible* properties, where the term "incompatible" has a very special quantum mechanical meaning with no exact classical counterpart.

• A quantum system cannot simultaneously possess two incompatible properties. For example, a spin-half particle cannot have both $S_x = 1/2$ and $S_z = 1/2$. There is nothing in the Hilbert space that could be used to represent such a combined property.

★ All the major conceptual difficulties of quantum theory are associated with the fact that the quantum Hilbert space allows incompatible properties.

• There is nothing analogous to this in classical physics, so knowing what to do (or not do) with incompatible properties is key to a clear understanding of quantum theory.

2 Operators

2.1 Definition

■ *Operators* are linear maps of the Hilbert space $\mathcal{H}$ onto itself. If A is an operator, then for any $|\psi\rangle$ in $\mathcal{H}$, $A|\psi\rangle$ is another element in $\mathcal{H}$, and linearity means that

$$A\big(b|\psi\rangle + c|\phi\rangle\big) = bA|\psi\rangle + cA|\phi\rangle \tag{17}$$

for any pair $|\psi\rangle$ and $|\phi\rangle$, and any two (complex) numbers b and c.

• The product AB of two operators A and B is defined by

$$(AB)|\psi\rangle = A\big(B|\psi\rangle\big) = AB|\psi\rangle, \tag{18}$$

where one usually omits the parentheses, as on the right side. Note that in general $AB \neq BA$.

2.2 Dyads and completeness

★ The simplest operator is a *dyad*, written in Dirac notation as a ket followed directly by a bra, e.g., $|\chi\rangle\langle\omega|$. Its action is defined by

$$\big(|\chi\rangle\langle\omega|\big)|\psi\rangle = |\chi\rangle\langle\omega|\psi\rangle = \big(\langle\omega|\psi\rangle\big)|\chi\rangle. \tag{19}$$

∘ The middle term is not really required for the definition, as the left side is defined by the right side: the scalar (complex number) $\langle\omega|\psi\rangle$ multiplying the ket $|\chi\rangle$. Nonetheless the middle term, formed by removing the parentheses and replacing two vertical bars $\|$ between ω and ψ with one bar $|$ is one of the examples of "Dirac magic" which makes this notation appealing to physicists.

• The following "completeness relation" is extremely useful:

$$I = \sum_j |j\rangle\langle j|. \tag{20}$$

Here I is the identity operator, $I|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle$, and the sum on the right is over dyads $|j\rangle\langle j|$ corresponding to the elements $|j\rangle$ of an orthonormal basis.

∘ Among the useful applications of (20):

$$|\psi\rangle = \big(\sum_j |j\rangle\langle j|\big)|\psi\rangle = \sum |j\rangle\langle j|\psi\rangle = \sum_j \langle j|\psi\rangle \cdot |j\rangle, \tag{21}$$

where a dot has been inserted in the final expression for clarity.

2.3 Matrices

★ Given an operator A and a basis $\{|\beta_j\rangle\}$, which need not be orthonormal, the *matrix* associated with A is the square array of numbers A_{jk} defined by:

$$A|\beta_k\rangle = \sum_j |\beta_j\rangle A_{jk} = \sum_j A_{jk}|\beta_j\rangle. \tag{22}$$

∘ The intermediate form with the matrix elements following the kets can provide a useful mnemonic for the order of the subscripts because it looks like the "natural" Dirac expression given below in (23).

∘ Note that the matrix depends on the choice of basis as well as on the operator A.

★ In the case of an orthonormal basis one can employ (20) to write

$$A|k\rangle = I \cdot A|k\rangle = \big(\sum_j |j\rangle\langle j|\big)A|k\rangle = \sum_j |j\rangle\langle j|A|k\rangle = \sum_j \langle j|A|k\rangle \cdot |j\rangle. \tag{23}$$

Here $\langle j|A|k\rangle$, the inner product of $|j\rangle$ with $A|k\rangle$, is the same as A_{jk} in (22). Thus $\langle j|A|k\rangle$, and more generally $\langle \psi|A|\omega\rangle$, is referred to as a "matrix element" when using Dirac notation.

• In a similar way

$$A = I \cdot A \cdot I = \sum_j |j\rangle\langle j| \cdot A \cdot \sum_k |k\rangle\langle k| = \sum_{jk} \langle j|A|k\rangle \cdot |j\rangle\langle k|. \tag{24}$$

allows us to express the operator A as a sum of dyads, with coefficients given by its matrix elements.

∘ The object written as $\langle j|A|k\rangle$ in Dirac notation corresponds to the matrix A_{jk} whose components are indicated using subscripts. Note that the order j before k is the same in both cases.

□ Exercise. Show that $\langle j|A|k\rangle = A_{jk}$, where A_{jk} is defined using (22) with $|\beta_j\rangle = |j\rangle$.

★ When A refers to a qubit the usual way of writing the matrix in the standard (or computational) basis is (note the order of the elements):

$$\begin{pmatrix} \langle 0|A|0\rangle & \langle 1|A|0\rangle \\ \langle 1|A|0\rangle & \langle 1|A|1\rangle \end{pmatrix} \tag{25}$$

• Another application of (20) is in writing the matrix element of the product of two operators in terms of the individual matrix elements:

$$\langle j|AB|k\rangle = \langle j|A \cdot I \cdot B|k\rangle = \sum_m \langle j|A|m\rangle\langle m|B|k\rangle. \tag{26}$$

Using subscripts this equation would be written as

$$\big(AB\big)_{jk} = \sum_m A_{jm} B_{mk}. \tag{27}$$

★ The *rank* of an operator A or its matrix is the maximum number of linearly independent rows of the matrix, which is the same as the maximum number of linearly independent columns.

• The rank does not depend upon which basis is used to produce the matrix for the operator.

• The rank of a dyad is 1.

★ The *trace* $\mathrm{Tr}(A)$ of an operator A is the sum of the diagonal elements of its matrix:

$$\mathrm{Tr}(A) = \sum_j \langle j|A|j\rangle = \sum_j A_{jj}. \tag{28}$$

One can show that the trace is independent of the basis used in define the matrix elements. In particular, one does not need to use an orthonormal basis; $\sum_j A_{jj}$ can be used with the A_{jj} defined in (22).

• The following are very useful formulas

$$\mathrm{Tr}\big(|\phi\rangle\langle\psi|\big) = \langle\psi|\phi\rangle, \quad \mathrm{Tr}\big(A|\phi\rangle\langle\psi|\big) = \langle\psi|A|\phi\rangle \tag{29}$$

□ Exercise. Derive the formulas in (29)

2.4 Dagger or adjoint

★ The dagger or adjoint operation † can be illustrated by some examples:

$$\big(|\psi\rangle\big)^\dagger = \langle\psi|, \quad \big(\langle\psi|\big)^\dagger = |\psi\rangle, \tag{30}$$

$$\big(b|\psi\rangle + c|\phi\rangle\big)^\dagger = b^*\langle\psi| + c^*\langle\phi|, \tag{31}$$

$$\big(|\psi\rangle\langle\omega|\big)^\dagger = |\omega\rangle\langle\psi| \tag{32}$$

$$\langle j|A^\dagger|k\rangle = \big(\langle k|A|j\rangle\big)^*, \tag{33}$$

$$\big(aA + bB\big)^\dagger = a^*A^\dagger + b^*B^\dagger, \tag{34}$$

$$(AB)^\dagger = B^\dagger A^\dagger. \tag{35}$$

○ Note that the dagger operation is *antilinear* in that scalars such as a and b are replaced by their complex conjugates. In fact, † can be thought of as a generalization of the idea of taking a complex conjugate, and in mathematics texts it is often denoted by *.

• The operator $A^\dagger$ is called the *adjoint* of the operator A. From (33) one sees that the matrix of $A^\dagger$ is the complex conjugate of the transpose of the matrix of A.

2.5 Normal operators

★ A *normal* operator A on a Hilbert space is one that commutes with its adjoint, $AA^\dagger = A^\dagger A$. Normal operators have the nice property that they can be diagonalized using an orthonormal basis, that is

$$A = \sum_j \alpha_j |a_j\rangle\langle a_j| \tag{36}$$

where the basis vectors $|a_j\rangle$ are *eigenvectors* or *eigenkets* of A and the (in general complex) numbers α_j are its *eigenvalues*. Equivalently, the matrix of A in this basis is diagonal

$$\langle a_j|A|a_k\rangle = \alpha_j\delta_{jk}. \tag{37}$$

• Equation (36) is often referred to as the *spectral form* or *spectral resolution* of the operator A.

∘ Note that the completeness relation (20) is of the form (36), since the eigenvalues of I are all equal to 1.

2.6 Hermitian operators

★ A *Hermitian* or *self-adjoint* operator A is defined by the property that $A = A^\dagger$, so it is a normal operator. It is the quantum analog of a real (as opposed to a complex) number. Its eigenvalues α_j are real numbers.

∘ The terms "Hermitian" and "self-adjoint" mean the same thing for a finite-dimensional Hilbert space, which is all we are concerned with; the distinction is important for infinite-dimensional spaces.

• Hermitian operators in quantum mechanics are used to represent *physical variables*, quantities such as energy, momentum, angular momentum, position, and the like. The operator representing the energy is the *Hamiltonian* H.

∘ The operator $S_z = \frac{1}{2}\big(|z^+\rangle\langle z^+| - |z^-\rangle\langle z^-|\big)$ represents the the z component of angular momentum (in units of $\hbar$) of a spin-half particle.

★ In classical physics a physical variable, such as the energy or a component of angular momentum, always has a well-defined value for a physical system in a particular state. In quantum physics this is no longer the case: if a quantum system is in the state $|\psi\rangle$, the physical variable corresponding to the operator A has a well-defined value if and only if $|\psi\rangle$ is an eigenvector of A, $A|\psi\rangle = \alpha|\psi\rangle$, where α, necessarily a real number since $A^\dagger = A$, is the value of the physical variable in this state.

∘ The eigenstates of S_z for a spin-half particle are $|z^+\rangle$ and $|z^-\rangle$, with eigenvalues of $+1/2$ and $-1/2$, respectively. Thus for such a particle the z component of angular momentum can take on only two values (it is "quantized"), in contrast to the (uncountably) infinite set of values available to a classical particle.

∘ If $|\psi\rangle$ is *not* an eigenstate of A, then in this state the physical quantity A is *undefined*, or *meaningless* in the sense that quantum theory can assign it no meaning.

∘ The state $|x^+\rangle$ is an eigenstate of S_x but not of S_z. Hence in this state S_x has a well-defined value (1/2), whereas S_z is undefined.

• There have been many attempts to assign a physical meaning to A when a quantum system is in a state which is not an eigenstate of A. All such attempts to make what is called a "hidden variable" theory have (thus far, at least) been unsuccessful.

∘ An apparent exception to this statement arises when "state" is used as a mathematical tool for calculating probabilities, rather than as an actual physical property of the system. We shall deal with this later when discussing quantum probabilities.

2.7 Projectors

★ A *projector*, short for "orthogonal projection operator", is a Hermitian operator $P = P^\dagger$ which is idempotent in the sense that $P^2 = P$. Equivalently, it is a Hermitian operator all or whose eigenvalues are either 0 or 1. Therefore there is always a basis (which depends, of course, on the projector) in which its matrix is diagonal in the sense of (37), with only 0 or 1 on the diagonal. Conversely, such a matrix always represents a projector.

• There is a one-to-one correspondence between a projector P and the *subspace* $\mathcal{P}$ of the Hilbert space that it projects onto. $\mathcal{P}$ consists of all the kets $|\psi\rangle$ such that $P|\psi\rangle = |\psi\rangle$; i.e., it is the eigenspace consisting of eigenvectors of P with eigenvalue 1.

∘ The term "projector" is used because such an operator "projects" a vector in a "perpendicular" manner onto a subspace. See Fig. 3.4 in CQT.

∘ Both the identity I and the zero operator 0 which maps every ket onto the zero ket are projectors.

∘ A more interesting example is the dyad $|\psi\rangle\langle\psi|$ for a normalized ($\|\psi\| = 1$) state $|\psi\rangle$. If $|\psi\rangle$ is not normalized (and not zero), the corresponding projector is

$$P = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} \tag{38}$$

∘ If $|\psi\rangle$ and $|\phi\rangle$ are two normalized states orthogonal to each other, $\langle\psi|\phi\rangle = 0$, then the sum $|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi|$ of the corresponding dyads is also a projector.

★ The physical significance of projectors is that they represent *physical properties* of a quantum system that can be either true or false. The property P corresponding to a projector P (it is convenient to use the same symbol for both) is true if the physical state $|\psi\rangle$ of the system is an eigenstate of P with eigenvalue 1, and false if it is an eigenstate with eigenvalue 0. If $|\psi\rangle$ is not an eigenstate of P, then the corresponding property is undefined (meaningless) for this state.

∘ For example, $|z^+\rangle\langle z^+|$, assuming $\langle z^+|z^+\rangle = 1$, is the projector for a spin half particle corresponding to the property $S_z = +1/2$. If the particle is in the state $|z^+\rangle$ the property is true, while if the particle is in the state $|z^-\rangle$ the property is false. In all other cases, such as the state $|x^+\rangle$, the property is undefined.

• The *negation* of a property P is represented by the projector $\tilde{P} = I - P$, also written as $\sim P$ or $\neg P$. If P is true, then $\tilde{P}$ is false, and vice versa.

• More generally, when two projectors P and Q are orthogonal, $PQ = 0$, the truth of one implies that the other is false. Note that $\tilde{P}P = (I - P)P = P - P^2 = P - P = 0$.

∘ The negation of $S_z = +1/2$ is $S_z = -1/2$, and vice versa.

★ Two quantum properties represented by projectors P and Q are said to be *compatible* if $PQ = QP$, i.e., if P and Q commute. Otherwise, when $PQ \neq QP$, they are *incompatible*.

• When $PQ = QP$, the product PQ is itself a projector, and represents the property "P AND Q," i.e, the property that the system has both properties P and Q at the same time. On the other hand it is impossible to make sense of the expression "P AND Q" when P and Q are incompatible. See the discussion in CQT Sec. 4.6. Attempting to combine incompatible properties violates the *single framework* rule of quantum interpretation, and leads sooner or later to contradictions and irresolvable paradoxes.

∘ The projectors $|z^+\rangle\langle z^+|$ and $|x^+\rangle\langle x^+|$ do not commute, and so $S_z = +1/2$ and $S_x = +1/2$ are examples of incompatible properties.

□ Exercise. Show that if P and Q are commuting projectors, then $P + Q - PQ$ is a projector. Argue that it represents "P OR Q" for the nonexclusive OR. What projector corresponds to the exclusive XOR?

2.8 Positive operators

★ The *positive* operators form another important class of Hermitian operators. They are defined by the requirement that the α_j in (36) be nonnegative, $\alpha_j \geq 0$, or equivalently by the requirement that for every ket $|\psi\rangle$

$$\langle\psi|A|\psi\rangle \geq 0. \tag{39}$$

∘ Both of these ways of characterizing a positive operator are useful for certain purposes, and both should be memorized.

□ Exercise. Show that these two definitions of a positive operator are equivalent in that each implies the other.

• Positive operators arise in quantum mechanics in various contexts, but one of the most important is when dealing with probabilities, which are inherently positive quantities.

2.9 Unitary operators

★ A unitary operator U has the property that

$$U^\dagger U = I = UU^\dagger. \tag{40}$$

• Since U commutes with its adjoint it is a normal operator and can be written in the form (36). Then (40) implies and is implied by the condition that all the eigenvalues of U are complex numbers of magnitude 1, i.e., they lie on the unit circle in the complex plane.

∘ In a finite-dimensional Hilbert space, with U mapping the space into itself, each equality in (40) implies the other, so that one need only check one of them, say $UU^\dagger = I$, to see if U is unitary.

∘ If one thinks of U as a matrix, the first equality in (40) is equivalent to the statement that the columns of U, thought of as column vectors, form an orthonormal basis of the Hilbert space. The second equality states that the rows of U likewise form an orthonormal basis.

□ Exercise. Show this.

• In quantum mechanics unitary operators are used to change from one orthonormal basis to another, to represent *symmetries*, such as rotational symmetry, and to describe some aspects of the *dynamics* or *time development* of a quantum system.

3 Bloch sphere

★ Any physical state $|\psi\rangle$ of a qubit (ray or normalized vector in the two-dimensional Hilbert space) can be associated in this way with a direction $w = (w_x, w_y, w_z)$ in space for which $S_w = 1/2$, i.e., the w component of angular momentum is positive. There is therefore a one-to-one correspondence between directions, or the corresponding points on the unit sphere, with rays of a two-dimensional Hilbert space. This is known as the *Bloch sphere representation* of qubit states.

• The usual correspondence employed in quantum information is that indicated in (4). Note that one can always multiply the ket by a nonzero complex number without changing the corresponding ray or the physical interpretation. Thus one will typically find (4) written in the form

$$\cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle \leftrightarrow S_w = +1/2, \quad \sin(\theta/2)|0\rangle - e^{i\phi}\cos(\theta/2)|1\rangle \leftrightarrow S_w = -1/2, \tag{41}$$

where the kets are normalized. (But even if the kets are normalized there is still a phase whose choice is arbitrary.)

∘ Note that it is the *surface* of the Bloch sphere—vectors w of unit length—that correspond to different rays. There is another role for the inside of the sphere, the "Bloch ball," which will come up later in a discussion of density operators.

★ Two states of a qubit are orthogonal, physically distinct or distinguishable, if they are antipodes, two points at opposite ends of a diagonal. For example $|0\rangle$ and $|1\rangle$ are the north and south pole. (Note the shorthand of $|0\rangle$ for the ray passing through $|0\rangle$. The north pole of the Bloch sphere also corresponds to $i|0\rangle$ or $2|0\rangle$.)

∘ Consequently, any orthonormal basis of a qubit is associated with a pair or antipodes of the Bloch sphere.

★ A linear operator maps a ray onto a ray, or onto a zero vector. Consequently, a linear operator on a qubit maps the Bloch sphere onto itself, or in the case of a noninvertible operator, onto a single point on the sphere.

• This map for a unitary operator corresponds to a proper rotation of the Bloch sphere.

∘ A proper rotation of a three-dimensional object is one that can be carried out physically, one that maps a right-handed object into its right-handed counterpart. Improper rotations map a right-handed object into its mirror image.

∘ A proper rotation can always be described in terms of a unit vector n denoting a direction in space and an angle ω (in radians) of rotation about n in the right-hand sense: with the thumb of your right hand in the direction n, your fingers point in the direction of positive rotation.

• Of particular importance are rotations of 180° about the x, y, and z axes, obtained using the unitary operators $X = \sigma_x$, $Y = \sigma_y$ and $Z = \sigma_z$, respectively. In the standard basis the corresponding matrices are the well-known Pauli matrices:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{42}$$

4 Composite systems and tensor products

★ In quantum theory the Hilbert space of a composite system (such as *two* qubits) is the *tensor product* of the Hilbert spaces for the subsystems.

∘ Composite systems, tensor products are discussed in CQT Ch. 6, QCQI Sec. 2.1.7

4.1 Definition

★ If two subsystems a and b that together comprise a total system, the Hilbert space of the latter is the *tensor product* of the Hilbert spaces of the subsystems, $\mathcal{H} = \mathcal{H}_a \otimes \mathcal{H}_b$. The dimension d of $\mathcal{H}$ is the product of the dimensions d_a and d_b of $\mathcal{H}_a$, $\mathcal{H}_b$. Let $\{|a_j\rangle\}$ and $\{|b_p\rangle\}$ be orthonormal bases of $\mathcal{H}_a$ and $\mathcal{H}_b$. The Hilbert space $\mathcal{H}$ consists of all vectors that can be written in the form

$$|\psi\rangle = \sum_j \sum_p M_{jp} \big(|a_j\rangle \otimes |b_p\rangle\big) \tag{43}$$

for some choice of the complex coefficients M_{jp}.

• Multiplication using $\otimes$ satisfies the usual distributive laws:

$$\big(|a\rangle + |a'\rangle\big) \otimes \big(|b\rangle + |b'\rangle\big) = |a\rangle \otimes |b\rangle + |a\rangle \otimes |b'\rangle + |a'\rangle \otimes |b\rangle + |a'\rangle \otimes |b'\rangle, \tag{44}$$

and scalar constants (complex numbers) can always be placed at the left:

$$\big(c|a\rangle\big) \otimes |b\rangle = |a\rangle \otimes \big(c|b\rangle\big) = c\big(|a\rangle \otimes |b\rangle\big) = c|a\rangle \otimes |b\rangle, \tag{45}$$

★ Bra vectors are formed from kets and vice versa using the dagger operation †:

$$\big(|a\rangle \otimes |b\rangle\big)^\dagger = \langle a| \otimes \langle b|; \quad \big(\langle a| \otimes \langle b|\big)^\dagger = |a\rangle \otimes |b\rangle \tag{46}$$

∘ Note that the order a to the left of b is not changed by the dagger operation; this is an exception to the "reverse the order" rule.

• Using (46) plus the fact that † is *anti*linear, one sees that if $|\psi\rangle$ is given by (43), then

$$\langle\psi| = \sum_j \sum_p M^*_{jp} \big(\langle a_j| \otimes \langle b_p|\big) \tag{47}$$

4.2 Product and entangled states

★ States in $\mathcal{H}$ of the form $|a\rangle \otimes |b\rangle$ are *product* states; all others are *entangled* states.

• Notation. $|a\rangle \otimes |b\rangle$ is often abbreviated to $|a\rangle|b\rangle$, or even to $|ab\rangle$ when the context makes plain what the symbols mean. Sometimes inserting a $\otimes$ is useful because it makes things clearer.

• It is often not immediately evident whether a state written in the form (43) is or is not a product state, for both $|a\rangle$ and $|b\rangle$ in the product state $|a\rangle \otimes |b\rangle$ may be linear combinations of a number of basis states $|a_j\rangle$ and $|b_p\rangle$. So it is useful to note that $|\psi\rangle$ in (43) is a product state if and only if the rank of the matrix M_{jp} is 1.

∘ Recall that the rank of a matrix is the number of linearly independent rows, which is always equal to the number of linearly independent columns.

★ Schmidt decomposition. Any $|\psi\rangle$ in $\mathcal{H}_a \otimes \mathcal{H}_b$ can be written as

$$|\psi\rangle = \sum_j \lambda_j |\hat{a}_j\rangle \otimes |\hat{b}_j\rangle, \tag{48}$$

where $\{|\hat{a}_j\rangle\}$ and $\{|\hat{b}_p\rangle\}$ are special orthonormal bases of $\mathcal{H}_a$ and $\mathcal{H}_b$ which depend upon the $|\psi\rangle$ one is considering. These bases can in addition be chosen so that the $\{\lambda_j\}$ are nonnegative real numbers, and for this reason (48) is sometimes written with $\sqrt{\lambda_j}$ in place of λ_j.

• The number of nonzero terms in (48) is equal to the rank of the matrix M_{jp} in (43), and is called the *Schmidt rank* of $|\psi\rangle$.

• A state is a product state if and only if its Schmidt rank is 1; entangled states always have a Schmidt rank of 2 or more.

★ Physical interpretation: A product state $|a\rangle \otimes |b\rangle$ has the meaning that system a has the property $|a\rangle$ (i.e., corresponding to the projector $|a\rangle\langle a|$) and system b the property $|b\rangle$. For an entangled state $|\psi\rangle$ one typically *cannot* assign definite properties to the individual systems a and b. However, we shall see later that if $|\psi\rangle$ is thought of not as representing an actual physical property, but as a *pre-probability*, then it can make sense to talk about probabilities about properties of the separate subsystems a and b.

• Consider the case of two qubits, and assume that

$$|\psi\rangle = \alpha|00\rangle + \beta|11\rangle, \tag{49}$$

with $\alpha \neq 0$, $\beta \neq 0$. One can show that the only projectors on $\mathcal{H}_a$ that commute with $P = |\psi\rangle\langle\psi|$ are I and 0, so in this case it is not possible to say that the composite system possesses property P and that the subsystem $\mathcal{H}_a$ possesses *any* nontrivial property.

4.3 Operators on tensor products

★ Product operators $A \otimes B$ act in the following way:

$$\big(A \otimes B\big)\big(|a\rangle \otimes |b\rangle\big) = \big(A|a\rangle\big) \otimes \big(B|b\rangle\big). \tag{50}$$

◦ One can use linearity and (43) to extend this to the action of $A \otimes B$ on any $|\psi\rangle$ in $\mathcal{H}_a \otimes \mathcal{H}_b$.

• Any operator on $A \otimes B$ can be written as a sum of product operators, so (50) suffices to define the action of any operator on any state of $\mathcal{H}_a \otimes \mathcal{H}_b$.

• A dyad constructed from two product states is a product operator. Note how one rearranges the terms to make this explicit:

$$\big(|a\rangle \otimes |b\rangle\big)\big(\langle a'| \otimes \langle b'|\big) = \big(|a\rangle\langle a'|\big) \otimes \big(|b\rangle\langle b'|\big) = |a\rangle\langle a'| \otimes |b\rangle\langle b'| \tag{51}$$

★ Adjoint: $\big(A \otimes B\big)^\dagger = A^\dagger \otimes B^\dagger$, and extend to other cases using the fact that † is antilinear. Note that A remains to the left of B after applying †.

★ Note that $A \otimes I$ is often written as A if it is obvious that the operator A refers to subsystem $\mathcal{H}_a$; similarly $I \otimes B$ is often written as B. This sometimes causes confusion.

★ Products of product operators:

$$(A \otimes B)(A' \otimes B') = AA' \otimes BB'. \tag{52}$$

4.4 Example of two qubits

★ We write product states $|a\rangle \otimes |b\rangle$ in the abbreviated form $|ab\rangle$. The computational (standard) basis of the two qubit system is formed by the four states

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle. \tag{53}$$

• Matrices (and column/row vectors) are written using these basis elements in (binary) numerical order, thus for an operator R:

$$\begin{pmatrix} \langle 00|R|00\rangle & \langle 00|R|01\rangle & \langle 00|R|10\rangle & \langle 00|R|11\rangle \\ \langle 01|R|00\rangle & \langle 01|R|01\rangle & \langle 01|R|10\rangle & \langle 01|R|11\rangle \\ \langle 10|R|00\rangle & \langle 10|R|01\rangle & \langle 10|R|10\rangle & \langle 10|R|11\rangle \\ \langle 11|R|00\rangle & \langle 11|R|01\rangle & \langle 11|R|10\rangle & \langle 11|R|11\rangle \end{pmatrix}. \tag{54}$$

◦ In the case of a product operator $R = A \otimes B$ one can think of the 4×4 matrix as consisting of four 2×2 blocks obtained by replicating the B matrix 4 times, and multiplying each by the corresponding matrix element of A.

★ The following entangled *Bell states* frequently arise in discussions of two qubits:

$$\begin{aligned} |B_0\rangle &= \big(|00\rangle + |11\rangle\big)/\sqrt{2}, \\ |B_1\rangle &= \big(|01\rangle + |10\rangle\big)/\sqrt{2}, \\ |B_2\rangle &= \big(|00\rangle - |11\rangle\big)/\sqrt{2}, \\ |B_3\rangle &= \big(|01\rangle - |10\rangle\big)/\sqrt{2}, \end{aligned} \tag{55}$$

◦ There is no standard notation for denoting a Bell state and sometimes it is convenient to use a different choice of phases. See p. 25 of QCQI for a slightly different notation.

• The states (55) form an orthonormal basis of the 2 qubit space.

• The state $|B_3\rangle$ is the spin singlet state used by Bohm in discussing the Einstein-Podolsky-Rosen paradox, and for this reason is sometimes called an “EPR” state, though that term is also sometimes used for other Bell states.

• Physical interpretation. One can think of $|B_0\rangle$ as “something like” a classical situation in which two bits, a and b are either both 0 with probability 1/2, or both 1 with probability 1/2. But there is no really good classical analogy for an entangled quantum state.

4.5 Multiple systems. Identical particles

★ Tensor products of three or more Hilbert spaces. For the most part these are obvious generalizations of the case of two spaces. Exception: There is no (satisfactory) generalization of the Schmidt decomposition (48) to three or more spaces.

★ The tensor product Hilbert space $\mathcal{H} = \mathcal{H}_a \otimes \mathcal{H}_b \otimes \mathcal{H}_c$ is equivalent to $(\mathcal{H}_a \otimes \mathcal{H}_b) \otimes \mathcal{H}_c$ or $\mathcal{H}_a \otimes (\mathcal{H}_b) \otimes \mathcal{H}_c$ or $(\mathcal{H}_a \otimes \mathcal{H}_c) \otimes \mathcal{H}_b$. It has an orthonormal basis $\{|a_j\rangle \otimes |b_p\rangle \otimes |c_s\rangle\}$ if $\{|a_j\rangle\}$, $\{|b_p\rangle\}$, and $\{|c_s\rangle\}$ are orthonormal bases of $\mathcal{H}_a$, $\mathcal{H}_b$, and $\mathcal{H}_c$, and $\mathcal{H}$ consists of all linear combinations of these basis vectors.

• The dimension d of $\mathcal{H} = \mathcal{H}_a \otimes \mathcal{H}_b \otimes \mathcal{H}_c$ is the product of the dimensions of the factors: $d = d_a \cdot d_b \cdot d_c$.

★ Identical particles. In quantum mechanics the tensor product space for identical particles is complicated, because of symmetry requirements. We will ignore these because usually we deal with quantum particles in separate locations. In other situations one can get away with treating the particles as if they were nonidentical by introducing fictitious “exchange interactions.”