qmd111

Hilbert Space Quantum Mechanics

Robert B. Griffiths
Version of 20 August 2012

Contents

References:

CQT = *Consistent Quantum Theory* by Griffiths (Cambridge, 2002). See in particular Ch. 2; Ch. 3; Ch. 4 except for Sec. 4.3; Ch. 5.

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 or spin half

★ 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*

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

(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 the quantum mechanics of atoms 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.

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

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

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

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

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

$$|w^+\rangle = \cos(\theta/2)|z^+\rangle + e^{i\phi}\sin(\theta/2)|z^-\rangle, \quad |w^-\rangle = \sin(\theta/2)|z^+\rangle - e^{i\phi}\cos(\theta/2)|z^-\rangle. \tag{4}$$

∘ The convention used in (3) and (4) is common but not universal. In particular, even if one supposes that the kets $|w^+\rangle$ and $|w^-\rangle$ are normalized, they can be multiplied by an arbitrary phase, a complex number of magnitude 1, without changing their physical significance. Thus, for example, a different convention for phases is employed in CQT Eq. (4.14). See comments in Sec. 1.5 below.

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 (but 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 $|z^+\rangle = |0\rangle$, 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

★ In quantum information theory 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.

• In atomic physics it is natural to think of $d = 3$ as "spin 1" and $d = 4$ as "spin 3/2", etc.

★ 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* $\{|b_j\rangle\}$, $j = 1, 2, \dots d$, with the property that

$$\langle b_j | b_k \rangle = \delta_{jk}. \tag{6}$$

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

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

where the coefficients v_j and w_j are given by

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

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

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

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

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

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 $\{|b_j\rangle\}$. 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 kets or two vectors $|\psi\rangle$ and $c|\psi\rangle$, where c is any *nonzero* complex number denote *exactly the same* physical property. For this reason it is sometimes helpful to say that the physical state corresponds not to a particular ket in the Hilbert space, but to the *ray*, or one-dimensional subspace, defined by the collection of all the complex multiples of a particular ket.

• 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 it is more convenient not to use normalized vectors, and even if normalization is desirable it can often be supplied at the end rather than in the middle of a calculation

• The state of a single qubit is always a linear combination of the basis vectors $|0\rangle$ and $|1\rangle$, or $|z^+\rangle$ and $|z^-\rangle$:

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

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

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

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

$$|0\rangle + \gamma|1\rangle = |z^+\rangle + \gamma|z^-\rangle \tag{16}$$

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 Probabilities

2.1 Sample spaces

★ Quantum mechanics differs from classical mechanics in that probabilities are an essential part of the quantum world.

★ Ordinary "classical" probability theory requires three things: a *sample space* $\mathcal{S}$, an *event algebra* $\mathcal{E}$, and a probability distribution $\mathcal{P}$. They are also needed in quantum mechanics.

★ Definition: A sample space is a *complete* set *mutually exclusive* possibilities.

• Example. A die (singular of dice) can land with s spots on the top surface, with $s = 1, 2, 3, 4, 5, 6$. These six possibilities are *mutually exclusive*—only one of them occurs if the die is rolled. Furthermore they include *all* possibilities that can occur. These six possibilities constitute the sample space.

• A finite sample space will suffice for the present discussion; infinite spaces are mathematically more complicated, but the essential ideas are the same.

★ We will use an *event algebra* that consists of all subsets of elements from the sample space, including the empty set and the sample space itself. Because the sample space determines the event algebra and vice versa it is safe to use the term *framework* to refer to either.

• Example: the die. Subsets include things like $\{1\}$, $\{3, 4\}$ or $\{2, 4, 6\}$. There are a total of $2^6 = 128$ distinct subsets if we include the empty set and the total sample space as "subsets".

★ Probabilities. To each element s of the sample space assign a real number $p_s \geq 0$ in such a way that the total is 1:

$$\sum_s p_s = 1. \tag{17}$$

• Intuitively one thinks of p_s is the fraction of cases in which s occurs if an experiment is repeated a large number of times.

∘ If $p_s = 1$, this s will always occur; if $p_s = 0$, this s will never occur.

∘ Note that we do *not* have to assume in the case of the die that every p_s is 1/6. This might be appropriate, but dice have sometimes shown up in Las Vegas that would be better described by choosing a different set of values. As long as the p_s are all nonnegative and sum to 1, that is all that is required in order to apply probability theory.

★ Probabilities are assigned to an element E of the event algebra using the formula

$$\Pr(E) = \sum_{s \in E} p_s. \tag{18}$$

Thus if $E = \{3, 5\}$, $\Pr(E) = p_3 + p_5$.

• If $E = \{s\}$ is a subset with just one element, it would be pedantic to write $\Pr(\{s\})$ in place of $\Pr(s) = p_s$.

★ When assigning probabilities to quantum systems it is absolutely essential to start with a properly defined *quantum sample space*, a complete collection of mutually exclusive quantum properties.

• Later, in Sec. 3.8, we will see how to do this in general. For present purposes it will suffice to employ some *orthonormal basis* $\{|b_j\rangle\}$ of the Hilbert space: a collection of kets which are mutually orthogonal and (for convenience) normalized, see Sec. 1.4.

• Since these kets are orthogonal they represent *distinct* physical properties, and because they form a basis for the Hilbert space it is impossible to add yet another ket corresponding to yet another distinct physical property, so this collection is complete.

★ Example. Qubit or spin half. The kets $|z^+\rangle$ and $|z^-\rangle$ are orthogonal, and form a basis of the two-dimensional Hilbert space. The corresponding properties $S_z = +1/2$ and $S_z = -1/2$ are mutually exclusive. One and only one can occur in any particular experiment (think of a silver atom traveling through a Stern-Gerlach apparatus) or on any particular occasion. Either $S_z = +1/2$ or $S_z = -1/2$, one or the other must be true, provided we are using the S_z framework in our discussion of probabilities.

• For such a simple sample space one has only one "adjustable" probability. For if we suppose the $\Pr(S_z = +1/2) = p$, then because the probabilities must sum to 1 it follows, (17), that $\Pr(S_z = -1/2) = 1-p$

★ Of course one can replace z with x in the foregoing discussion. The S_x framework consists of the events $S_x = +1/2$ and $S_x = -1/2$. If one supposes that $\Pr(S_x = +1/2) = q$, then necessarily $\Pr(S_x = -1/2) = 1-q$.

• But the ket $|x^+\rangle$ is not a multiple of either $|z^+\rangle$ or $|z^-\rangle$, nor is it orthogonal to either of them. Hence the property $S_x = +1/2$ is *incompatible* with either of the properties $|z^+\rangle$ or $|z^-\rangle$. There is no way of including $|x^+\rangle$ in the S_z framework. Discussions involving $|x^+\rangle$ have to be kept strictly separate from discussions involving $|z^+\rangle$ or $|z^-\rangle$. This is an example of a fundamental principle of quantum mechanics known as the *single framework rule*.

!! In particular the pair $\{|x^+\rangle, |z^-\rangle\}$ does *not* constitute a sample space for a spin half particle.

☐ Exercise. For a spin half particle some of the following collections form a sample space and some do not. Give reasons in each case
(i) $\{|y^+\rangle\}$
(ii) $\{|y^+\rangle, |y^-\rangle\}$
(iii) $\{|z^-\rangle, |z^+\rangle, |y^-\rangle\}$
(iv) $\{|z^-\rangle, |y^+\rangle\}$

2.2 Born Rule

★ Let the orthonormal basis $\{|b_j\rangle\}$ be the quantum sample space we are interested in. The event algebra consists, just as in ordinary classical probability theory, of all possible subsets of the elements of the sample space. Probabilities for these events are determined once probabilities, a collection of nonnegative numbers $\{p_j\}$, with p_j the probability of $|s_j\rangle$, have been assigned to the elements of the sample space.

• How do we decide on the $\{p_j\}$? This is a complicated question, and, just as in ordinary classical applications of probability theory there is no single rule that will cover all cases. Sometimes probabilities are assigned by guesswork, sometimes by using the results of experiments to fit parameters, etc.

★ However, there is a uniquely quantum mechanical way of assigning probabilities, with no classical analog. It is known as the *Born rule*. We introduce it here simply as a calculational device. Later on we will see a variety of physical situations to which it can be applied, and that will add some physical intuition to the mathematical rule.

★ Let $|\psi\rangle$ be a normalized ket, $\langle\psi|\psi\rangle = 1$. We shall call it a *pre-probability*, because it can is used to generate a probability distribution. Let $\{|b_j\rangle\}$ be an orthonormal basis, the sample space we are interested in. The Born rule assigns probability

$$p_j = |\langle b_j|\psi\rangle|^2 = |\langle\psi|b_j\rangle|^2 \tag{19}$$

to the ket $|b_j\rangle$.

◦ If $|\psi\rangle$ is not normalized it can still be used as a pre-probability; in place of (19) use

$$p_j = |\langle\psi|b_j\rangle|^2/\langle\psi|\psi\rangle. \tag{20}$$

Or replace $|\psi\rangle$ in (19) with the normalized $|\bar{\psi}\rangle = |\psi\rangle/\sqrt{\langle\psi|\psi\rangle}$.

★ It is important to distinguish a *pre-probability* from a quantum *property*. Both of them are represented by kets—later we will generalize both concepts—so it is easy to confuse them. The fundamental difference is that a property is something real, whereas a pre-probability is an abstraction, just like an ordinary probability.

◦ To be sure, a vector in the Hilbert space is itself a mathematical abstraction, not physical reality. It can, however, *represent* a physical property in much the same sense that three numbers in an appropriate coordinate system can represent the present position of the center of mass of the planet Jupiter. A probability cannot correspond to a physical property in the same way (unless the probability is 0 or 1, in which case it becomes a certainty).

★ Example. Let $|\psi\rangle = |x^+\rangle$ be the pre-probability, and $\{|z+\rangle, |z^-\rangle\}$ the sample space. Then since $|\langle x^+|z^+\rangle| = |\langle x^+|z^-\rangle| = 1/\sqrt{2}$, we conclude that $\Pr(z^+) = \Pr(z^-) = 1/2$. If, on the other hand, we use the same pre-probability but a different sample space $\{|x^+\rangle, |x^-\rangle\}$, then $\Pr(x^+) = 1$, $\Pr(x^-) = 0$.

!! Beware! In the preceding example one may be tempted to combine the results and say that: "the probability to x^+ is 1, the probabilities of z^+ and z^- are 1/2, and the probability of x^- is 0." The trouble with this statement is that it suggests that $|x^+\rangle$, $|z^+\rangle$, $|z^-\rangle$, and $|x^-\rangle$ are all elements of a single event algebra, and one is comparing them using a common probability distribution. But $|x^+\rangle$ and $|z^+\rangle$ are incompatible (in the quantum sense), and it makes no sense to compare their probabilities.

2.3 Measurements

★ The Born rule is often stated in quantum textbooks in the following way, where to make it definite we consider the case of a spin half particle. Suppose that the normalized quantum state is $|\psi\rangle$. Then one finds statements such as:

M. "If S_z is measured the probability of finding $S_z = 1/2$ (in units of $\hbar$) is $|\langle z^+|\psi\rangle|^2$, and the probability of finding $S_z = -1/2$ is $|\langle z^-|\psi\rangle|^2$."

• This can be interpreted in the following way. Mentioning a particular physical variable, in this case S_z, implicitly defines a quantum sample space $\{|z^+\rangle, |z^-\rangle\}$ (or the corresponding decomposition of the identity, Sec. 3.8). Next comes the notion of an *ideal measurement* in which the measurement outcome is indicated on the measuring instrument by clearly distinguished macroscopic ("classical") states of affairs. In a tradition that goes back much earlier than modern electronics one speaks of the different outcomes as distinct positions of a *pointer*. An ideal measurement is one in which the apparatus pointer position accurately reflects the state of affairs that existed just before the measurement was carried out. The probabilities referred to in M are then probabilities that the measurement pointer will be in one position or the other.

★ There are two reasons for this circumlocution with its reference to "measurements".

• The first is tradition: the textbook writer learned quantum mechanics from a previous book written by someone who learned quantum mechanics from a previous book written by . . . by John von Neumann with the title *Mathematische Grundlagen der Quantenmechanik*. Von Neumann was a bit confused about this aspect of quantum theory, and his confusion has been passed down to his intellectual grandchildren.

• Second, measurements are being used to cover up conceptual difficulties of the quantum world. Hilbert space quantum mechanics allows one to speak in a sensible way about S_x for a spin half particle, or about S_z, or any other component, but combining descriptions of angular momentum components referring to different directions in space is not possible. Rather than facing up to this difficulty directly what the textbooks state is that there is impossible to simultaneously measure both S_x and S_z for a spin half particle.

• This statement, that S_x and S_z cannot be simultaneously measured, is correct. What textbooks omit is the explanation: what does not exist cannot be measured! There just is nothing in the quantum Hilbert space of a spin half particle that could represent the particle having both a value (necessarily $\pm 1/2$) of S_x *and* a value of S_z. Even a very competent experimentalist cannot measure what is not there! (Indeed, this is one thing that distinguishes them from less competent colleagues.)

★ Later on we will discuss the measurement process, using a simplified but fully consistent quantum mechanical description that includes the apparatus as well as the measured system.

• Until then, the reader should interpret textbook references to "measurement" as an indirect way of identifying a quantum sample space,

∘ Textbook statement: "Given that the quantum state is $|\psi\rangle$, what is the probability that the measurement of S_z will yield the value of $+1/2$?

∘ Translation: "Given the quantum state $|\psi\rangle$ understood as a pre-probability, and that the sample space consists of $|z^+\rangle$ and $|z^-\rangle$, what is the probability of $|z^+\rangle$?

3 Operators

3.1 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

$$A\big(b|\psi\rangle + c|\phi\rangle\big) = bA|\psi\rangle + cA|\phi\rangle \tag{21}$$

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

$$(AB)|\psi\rangle = A\big(B|\psi\rangle\big) = AB|\psi\rangle, \tag{22}$$

where one usually omits the parentheses, as on the right side.

• Note that in general $AB \neq BA$, the product of two operators depends upon the order. If $AB = BA$ one says that the operators A and B *commute* with each other; otherwise they do not commute.

3.2 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

$$\big(|\chi\rangle\langle\omega|\big)|\psi\rangle = |\chi\rangle\langle\omega|\psi\rangle = \big(\langle\omega|\psi\rangle\big)|\chi\rangle. \tag{23}$$

∘ 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", where $\{|b_j\rangle\}$ is any orthonormal basis, is extremely useful:

$$I = \sum_j |b_j\rangle\langle b_j|. \tag{24}$$

Here I is the identity operator, $I|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle$, and the sum on the right is over the dyads $|b_j\rangle\langle b_j|$ formed from the elements of the orthonormal basis.

∘ Among the useful applications of (24):

$$|\psi\rangle = \Big(\sum_j |b_j\rangle\langle b_j|\Big)|\psi\rangle = \sum |b_j\rangle\langle b_j|\psi\rangle = \sum_j \langle b_j|\psi\rangle \cdot |b_j\rangle, \tag{25}$$

where a dot has been inserted in the final expression for clarity. Thus one has a good mnemonic for the expansion coefficients of an arbitrary ket in some orthonormal basis.

3.3 Matrices

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

$$A|\beta_k\rangle = \sum_j |\beta_j\rangle A_{jk} = \sum_j A_{jk}|\beta_j\rangle. \tag{26}$$

∘ 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 (27).

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

★ In the case of an orthonormal basis $\{|b_j\rangle\}$, where for convenience we will use $|j\rangle$ as an abbreviation for $\{|b_j\rangle\}$, one can employ the completeness relation (24) to write

$$A|k\rangle = I \cdot A|k\rangle = \Big(\sum_j |j\rangle\langle j|\Big) A|k\rangle = \sum_j |j\rangle\langle j|A|k\rangle = \sum_j \langle j|A|k\rangle \cdot |j\rangle. \tag{27}$$

Here $\langle j|A|k\rangle$, the inner product of $|j\rangle$ with $A|k\rangle$, is the same as A_{jk} in (26). Note that the order j before k is the same in both cases. Thus $\langle j|A|k\rangle$ is referred to as a "matrix element" when using Dirac notation.

∘ Indeed, anything of the form $\langle\psi|A|\omega\rangle$ is referred to as a "matrix element", even when no matrix is in view!

• In a similar way

$$A = I \cdot A \cdot I = \sum_j |j\rangle\langle j| \cdot A \cdot \sum_k |k\rangle\langle k| = \sum_{jk} \langle j|A|k\rangle \cdot |j\rangle\langle k|. \tag{28}$$

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

☐ Exercise. Show that $\langle j|A|k\rangle = A_{jk}$, where A_{jk} is defined using (26) with $|\beta_j\rangle = |j\rangle$.

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

$$\begin{pmatrix} \langle 0|A|0\rangle & \langle 0|A|1\rangle \\ \langle 1|A|0\rangle & \langle 1|A|1\rangle \end{pmatrix} = \begin{pmatrix} \langle z^+|A|z^+\rangle & \langle z^+|A|z^-\rangle \\ \langle z^-|A|z^+\rangle & \langle z^-|A|z^-\rangle \end{pmatrix} \tag{29}$$

• Another application of (24) is in writing the matrix element of the product of two operators in terms of the individual matrix elements:

$$\langle j|AB|k\rangle = \langle j|A \cdot I \cdot B|k\rangle = \sum_m \langle j|A|m\rangle\langle m|B|k\rangle. \tag{30}$$

Using subscripts this equation would be written as

$$\big(AB\big)_{jk} = \sum_m A_{jm} B_{mk}. \tag{31}$$

★ 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:

$$\mathrm{Tr}(A) = \sum_j \langle j|A|j\rangle = \sum_j A_{jj}. \tag{32}$$

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 (26).

• The following are very useful formulas

$$\mathrm{Tr}\big(|\phi\rangle\langle\psi|\big) = \langle\psi|\phi\rangle, \quad \mathrm{Tr}\big(A|\phi\rangle\langle\psi|\big) = \langle\psi|A|\phi\rangle \tag{33}$$

☐ Exercise. Derive the formulas in (33)

3.4 Dagger or adjoint

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

$$\big(|\psi\rangle\big)^\dagger = \langle\psi|, \quad \big(\langle\psi|\big)^\dagger = |\psi\rangle, \tag{34}$$

$$\big(b|\psi\rangle + c|\phi\rangle\big)^\dagger = b^*\langle\psi| + c^*\langle\phi|, \tag{35}$$

$$\big(|\psi\rangle\langle\omega|\big)^\dagger = |\omega\rangle\langle\psi| \tag{36}$$

$$\langle j|A^\dagger|k\rangle = \big(\langle k|A|j\rangle\big)^*, \tag{37}$$

$$\big(aA + bB\big)^\dagger = a^*A^\dagger + b^*B^\dagger, \tag{38}$$

$$(AB)^\dagger = B^\dagger A^\dagger. \tag{39}$$

◦ 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 (37) one sees that the matrix of $A^\dagger$ is the complex conjugate of the transpose of the matrix of A.

3.5 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

$$A = \sum_j \alpha_j |a_j\rangle\langle a_j| \tag{40}$$

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

$$\langle a_j|A|a_k\rangle = \alpha_j\delta_{jk}. \tag{41}$$

• Equation (40) is often referred to as the *spectral form* or *spectral resolution* of the operator A.

◦ Note that the completeness relation (24) is of the form (40), since the eigenvalues of I are all equal to 1.

3.6 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.

★ However a ket $|\psi\rangle$ which is not an eigenstate of A can be employed as a *pre-probability* to assign probabilities to the different values of A, as discussed in Sec. 3.8 below.

3.7 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 (41), 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 $|\psi\rangle$, for which it is convenient to use the abbreviation

$$[\psi] = |\psi\rangle\langle\psi| \tag{42}$$

If $|\psi\rangle$ is not normalized (and not zero), the corresponding projector is $|\psi\rangle\langle\psi|$ divided by $\langle\psi|\psi\rangle$.

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

☐ Exercise. Prove it.

★ 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, $|z^+\rangle\langle z^+|$ 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?

3.8 Decomposition of the identity

★ A very important concept in Hilbert space quantum mechanics is that of a *decomposition of the identity*. This is a collection $\{P_j\}$ of projectors that sum to the identity:

$$I = \sum_j P_j, \quad P_j = P_j^\dagger = P_j^2, \quad P_j P_k = \delta_{jk} P_j. \tag{43}$$

Here the equation $P_j = P_j^\dagger = P_j^2$ has been included just to remind us that we are dealing with projectors. The final equation, $P_j P_k = \delta_{jk} P_j$, which says that the different projectors are orthogonal to each other, is a (nontrivial) consequence of the fact that $\sum_j P_j = I$.

∘ The term *projective decomposition of the identity* is sometimes used to distinguish what we are talking about from a POVM (positive operator-valued measure), see Sec. 3.9 below.

★ By a *quantum sample space* we shall always mean some decomposition of the identity of the form (43). The different projectors in the decomposition represent a collection of distinct quantum properties which are mutually exclusive—this is the significance of $P_j P_k = 0$ for $j \neq k$—and which is complete in the sense that these projectors sum to the identity operator I, which represents the property which is always true.

• Given such a sample space a pre-probability $|\psi\rangle$ (assumed normalized) can be used to assign probabilities to the different projectors, or the different subspaces onto which they project, using the formula

$$p_j = \langle\psi|P_j|\psi\rangle, \tag{44}$$

which generalizes the Born rule discussed earlier in Sec. 2.2

□ Exercise. Show that if $|\psi\rangle$ in (44) is normalized, $p_j \geq 0$ and $\sum_j p_j = 1$.

□ Exercise. Show that the right side of (44) is equal to $\mathrm{Tr}([\psi]P_j)$, that is, the trace of product of the two projectors $[\psi]$ and P_j; $[\psi]$ is defined in (42).

□ Exercise. Show that if the decomposition $\{P_j\}$ corresponds to an orthonormal basis, (44) is identical to (19).

★ An alternative way of writing the spectral form (40) of a normal operator is:

$$A = \sum_j \alpha'_j P_j; \quad \alpha'_j \neq \alpha'_k \text{ for } j \neq k. \tag{45}$$

• The difference between (40) and (45) is the following. It is possible that some eigenvalues occur repeatedly in the sum (40). It so, then these terms are collected together and the sum of all the dyads associated with a single eigenvalue forms one of the projectors P_j in (45). Here is a simple example to illustrate the idea:

$$A = \alpha'_1\bigl(|a_1\rangle\langle a_1| + |a_2\rangle\langle a_2|\bigr) + \alpha'_2\bigl(|a_3\rangle\langle a_3| + |a_4\rangle\langle a_4|\bigr), \tag{46}$$

where $\alpha'_1 \neq \alpha'_2$. In the notation of (45) we would set $P_1 = |a_1\rangle\langle a_1| + |a_2\rangle\langle a_2|$ and $P_2 = |a_3\rangle\langle a_3| + |a_4\rangle\langle a_4|$, and in the notation of (40) we would have $\alpha_1 = \alpha_2 = \alpha'_1$, $\alpha_3 = \alpha_4 = \alpha'_2$. Thus the use of primes on the eigenvalues in (45) is simply to avoid confusion in the use of subscripts; the actual values are the same.

★ Note that the spectral form (45) associates a *unique* decomposition of the identity with any Hermitian operator A: the projectors correspond to different subspaces of the Hilbert space corresponding to the different values, i.e., eigenvalues, this operator can take on.

• Consequently, if there is a scheme for assigning probabilities to these subspaces—e.g., using a pre-probability as in (44)— one can regard these as the probabilities that the the observable takes on a particular value:

$$\Pr(A = \alpha'_j) = p_j \tag{47}$$

○ In particular, if an ideal measurement of the quantity A is carried out, p_j is the probability that the measurement outcome (pointer position) will correspond to α'_j.

3.9 Positive operators

★ The *positive* operators form another important class of Hermitian operators. They are defined by the requirement that the α_j in (40) be nonnegative, $\alpha_j \geq 0$, or equivalently by the requirement that for every ket $|\psi\rangle$

$$\langle\psi|A|\psi\rangle \geq 0. \tag{48}$$

○ 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. In particular a POVM (positive operator-valued measure) is a collection $\{A_j\}$ of positive operators that sum to the identity:

$$\sum_j A_j = I. \tag{49}$$

The (projective) decomposition of the identity introduced earlier in (43) is an example of a POVM.

○ POVMs are used all the time in quantum information theory, and are useful for discussing non-ideal quantum measurements.

3.10 Unitary operators

★ A unitary operator U has the property that

$$U^\dagger U = I = UU^\dagger. \tag{50}$$

• Since U commutes with its adjoint it is a normal operator and can be written in the form (40). Then (50) 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 (50) 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 (50) 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.

4 Bloch Sphere

★ Any state $|\psi\rangle$ of a spin half particle or qubit regarded as a physical property, and thus associated with a ray in the two-dimensional Hilbert space, can be associated with a direction $w = (w_x, w_y, w_z)$ in ordinary three-dimensional space for which $S_w = 1/2$, i.e., the w component of angular momentum is positive, in the same way that $|z^+\rangle$ corresponds to a ray for which $S_z = +1/2$. There is thus 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*, and it is an extremely convenient way of thinking about such states. .

• The usual correspondence employed in quantum information and in atomic physics is that indicated in (4), though sometimes different overall phases are used. (The choice of a phase does not alter the significance of the ket as a quantum property, nor does it change the probabilities obtained when the ket is used as a pre-probability.)

∘ 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 $|z^+\rangle$ ($|0\rangle$) and $|z^-\rangle$ ($|1\rangle$) 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.

!! Normally one thinks of "orthogonal" as corresponding to "perpendicular," However, in the Bloch sphere picture the angle between two unit vectors proceeding from the center to (quantum mechanically) orthogonal states is 180°, not 90°. This can sometimes be confusing.

★ 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, see (29), the corresponding matrices are the well-known Pauli matrices:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{51}$$