PH 500 Problem Set #1

In this problem set, you'll review a variety of results from linear algebra, introducing the notation and extension to complex numbers we'll need in quantum mechanics.

I. VECTORSPACES IN DIRAC NOTATION

In a modern approach to quantum mechanics, the possible quantum states of a system are described as *vectors*. They're not ordinary three-component vectors, but that case will serve as a useful analogy. So we'll start by reviewing the basic rules for manipulating vectors. We'll use *Dirac notation*, which at first will seem bizarre and unfamiliar, but eventually will prove extremely convenient. In this notation, we'll write a quantum "state vector" as $|\psi\rangle$. The set of *all* quantum states forms a vectorspace ("vector" and "state" are synonymous here; for some reason people don't usually use the term "statespace," although it would be perfectly sensible).

A. Vectorspaces and Bases

A vectorspace is a set of states (vectors) $|\psi\rangle$ with the following properties:

- We can add states, $|\psi_1\rangle + |\psi_2\rangle$, and obtain another state in the vectorspace.
- We can multiply states by a complex constant, $c|\psi\rangle$, and obtain another state in the vectorspace. (Technically, we are considering a vectorspace over the field of complex numbers.)
- These operations are associative, commutative, and distributive in the usual way.
- There is an additive identity $|\text{zero}\rangle$ such that $|\psi\rangle + |\text{zero}\rangle = |\psi\rangle$. (This state is usually just written as 0, which is convenient but a bit sloppy, like writing $\vec{v} = 0$ instead of $\vec{v} = \vec{0}$; it would be more logical to to write this state as $|0\rangle$, but that notation is often reserved for the *ground state* of a system, which is not the same thing.) There is also a multiplicative identity 1 such that $1|\psi\rangle = |\psi\rangle$ and a multiplicative zero 0 such that $0|\psi\rangle = |\text{zero}\rangle$.
- Every $|\psi\rangle$ has an additive inverse $-|\psi\rangle$, such that $|\psi\rangle + (-|\psi\rangle) = |\text{zero}\rangle$, and every c except $c = 0$ has a multiplicative inverse $1/c$ such that $c \cdot (1/c) = 1$.

Of course, these are just the usual manipulations you are accustomed to when working with vectors.

As you know, we usually describe vectors by their *coordinates*. In ordinary three-dimensional space, this means writing an arbitrary vector $\boldsymbol{v}$ in terms of $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$. We'd like to generalize this idea to our vectorspace of quantum states. Let's consider a subset of our vectorspace consisting of the states $\{|e_1\rangle,\ |e_2\rangle,\ \ldots\}$. I'm intentionally being a little vague about how many states are in this subset, because in many cases we are interested in, there will will be infinitely many (instead of three as in the usual case). We say that this subset is *linearly independent* if there is no way to write any one element of the subset as a linear combination of the others. Equivalently, the subset is linearly independent if the only way to satisfy the equation

$$0 = c_1|e_1\rangle + c_2|e_2\rangle + \ldots \tag{1}$$

is for $c_1 = c_2 = \ldots = 0$. (This definition is equivalent to the previous one because if it was possible to satisfy this equation with nonzero coefficients, we could solve for one of the vectors with a nonzero coefficient in terms of the others.) Roughly, a linearly independent subset of the vectorspace doesn't have "too many" vectors. On the other hand, we say that our subset *spans* the space if every vector $|\psi\rangle$ can be written as a linear combination of elements of the subset,

$$\psi = c_1|e_1\rangle + c_2|e_2\rangle + \ldots , \tag{2}$$

for some choice of $c_1, c_2, \ldots$. Roughly, a subset that spans the vectorspace has "enough" vectors.

If a subset is both linearly independent and spans the vectorspace, we call it a *basis*. The *dimension* of the vectorspace is the number of elements in this subset. Although there can be many different bases for the same vectorspace, one can prove that they all have the same number of elements, so that the dimension is a property of the space, independent of the particular basis we choose. (It is perfectly possible, however, for the dimension to be *infinite*; but then all bases will have infinitely many elements.)

Since a basis spans the space, we can write any vector in the form of Eq. (2). Then $c_1, c_2, \ldots$ are the *coordinates* of the state $|\psi\rangle$ in this basis. Because the basis is linearly independent, these coordinates are *unique* — there is no other combination that would work. In more traditional notation, the vector is represented by its coordinates as a column, $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}$. We won't use that notation here, because we want to avoid committing to a particular basis.

B. Linear Operators

We have seen how to form new states as linear combinations — sums and scalar products — of other states. We will also be interested in *operators*, which are functions that take a state and return another state. For an operator $\hat{A}$, we'll write the result of this operation as $\hat{A}|\psi\rangle$. In particular, in quantum mechanics we will almost exclusively be interested in *linear* operators, for which $\hat{A}(|\psi_1\rangle + |\psi_2\rangle) = \hat{A}|\psi_1\rangle + \hat{A}|\psi_2\rangle$ and $\hat{A}(c|\psi\rangle) = c\hat{A}|\psi\rangle$. Note that this is a *very* restrictive definition. For example, we often refer to a function $y = mx + b$ as "linear," but our definition of linear only admits the case $b = 0$ (for $b \neq 0$ it would be an "affine" function).

We can fully describe a linear operator by specifying how it acts on a basis, since by linearity we can figure out what it does to any other state by taking linear combinations. Furthermore, we can decompose the result of acting on a particular basis element — which is another state — in that same basis. So we have

$$\begin{aligned} \hat{A}|e_1\rangle &= A_{11}|e_1\rangle + A_{21}|e_2\rangle + \ldots \\ \hat{A}|e_2\rangle &= A_{12}|e_1\rangle + A_{22}|e_2\rangle + \ldots \\ \vdots \quad & \quad \vdots \quad \quad \vdots \end{aligned} \tag{3}$$

and we can completely specify the linear operator by the complex numbers A_{ij}. In traditional

notation we would write these numbers as a matrix, $\begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$, but again, we generally won't use this notation, because we want to avoid committing to a particular basis.

C. Inner Product

The last operation we would like to define for our generalized vectorspace of quantum states is the *inner product*, which is the generalization of the *dot product* of ordinary vectors. (We won't generalize the *cross* product, which in its usual form is specific to three dimensions.) Let's review the situation in that case. From two vectors we obtain a number, which gives us information about how closely aligned the two vectors are. We start by writing two vectors in the standard basis,

$$\boldsymbol{v} = v_x\hat{\boldsymbol{x}} + v_y\hat{\boldsymbol{y}} + v_z\hat{\boldsymbol{z}} \quad \text{and} \quad \boldsymbol{w} = w_x\hat{\boldsymbol{x}} + w_y\hat{\boldsymbol{y}} + w_z\hat{\boldsymbol{z}}\,, \tag{4}$$

or, more formally,

$$\boldsymbol{v} = \sum_{j=1}^{3} v_j\boldsymbol{e}_j \quad \text{and} \quad \boldsymbol{w} = \sum_{j=1}^{3} w_j\boldsymbol{e}_j \tag{5}$$

where $\boldsymbol{e}_1 = \hat{\boldsymbol{x}}$, $\boldsymbol{e}_2 = \hat{\boldsymbol{y}}$, and $\boldsymbol{e}_3 = \hat{\boldsymbol{z}}$. Then their inner product is

$$\boldsymbol{v}\cdot\boldsymbol{w} = v_x^*w_x + v_y^*w_y + v_z^*w_z = \sum_{j=1}^{3} v_j^*w_j = |\boldsymbol{v}||\boldsymbol{w}|\cos\theta\,, \tag{6}$$

where θ is the angle between the two vectors. This formula may not look exactly like the one you're familiar with, because of the conjugation symbol on v_j. However, I claim it is the only sensible definition when we are considering *complex* vectors (that is, vectors with complex coordinates). Why? Well, for the special case when the two vectors are the *same*, we want the dot product to give a definition of length:

$$\boldsymbol{v}\cdot\boldsymbol{v} = \sum_{j=1}^{3} v_j^*v_j = |\boldsymbol{v}|^2\,. \tag{7}$$

With the conjugation, this is the sum of the magnitude squared (in the sense of complex numbers) of the three coordinates, which is a sensible definition of the total magnitude squared (in the sense of vectors). Without the conjugation, it could be negative or imaginary! Note, however, that this change means that the inner product is *no longer commutative*: $\boldsymbol{v}\cdot\boldsymbol{w} = (\boldsymbol{w}\cdot\boldsymbol{v})^*$.

To implement the inner product for our more abstract quantum states, it is helpful to make one more definition. For any state $|\psi\rangle$, we associate a "dual" state $\langle\psi|$. (The notation was designed for this purpose!). We furthermore stipulate that for any complex coefficient c, the state $c|\psi\rangle$ is associated with the dual state $\langle\psi|c^*$. The order here is not important (we can move a number around however we want), but the conjugate is — it will allow us to keep track of the conjugates introduced by the inner product. As we'd expect, the dual of the sum of two vectors is simply the sum of their duals. In traditional coordinate notation, then, the dual state is a *row* vector, so the

dual of $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}$ is $\begin{pmatrix} c_1^* & c_2^* & \cdots \end{pmatrix}$, but again we won't tend to use this notation because it commits us to a particular basis.

We then define the inner product to pair a state with a dual state:

$$[\text{inner product of } \phi \text{ and } \psi] = \langle \phi | \psi \rangle \,, \tag{8}$$

which will keep track of the conjugates automatically. Let's see how this works in terms of coordinates. First, let's go back to the case of ordinary vectors. To be completely precise, we should write

$$\boldsymbol{v} \cdot \boldsymbol{w} = \left(\sum_{i=1}^{3} v_i^* \hat{\boldsymbol{e}}_i \right) \cdot \left(\sum_{j=1}^{3} w_j \hat{\boldsymbol{e}}_j \right) \,. \tag{9}$$

We then use the fact that the $\hat{\boldsymbol{e}}_j$ are *orthonormal*: each $\hat{\boldsymbol{e}}_j$ is perpendicular to all the others and has length 1, so that

$$\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \,. \end{cases} \tag{10}$$

Even though there's nothing particularly complicated about the right-hand side of this expression, it arises so commonly that physicists define it as the Kronecker δ symbol,

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \,. \end{cases} \tag{11}$$

So in our quantum state notation, suppose $\{|e_j\rangle\}$ is an orthonormal basis, meaning it is a basis such that

$$\langle e_i | e_j \rangle = \delta_{ij} \,. \tag{12}$$

If we express two states in this basis,

$$\begin{aligned} |\psi\rangle &= a_1 |e_1\rangle + a_2 |e_2\rangle + \ldots = \sum_i a_i |e_i\rangle \\ |\phi\rangle &= b_1 |e_1\rangle + b_2 |e_2\rangle + \ldots = \sum_j b_j |e_j\rangle \,, \end{aligned} \tag{13}$$

then to take their inner product we first find the dual of $|\phi\rangle$,

$$\langle \phi | = \langle e_1 | b_1^* + \langle e_2 | b_2^* + \ldots = \sum_j \langle e_j | b_j^* \,, \tag{14}$$

and compute the inner product as

$$\langle \phi | \psi \rangle = \sum_j \langle e_j | b_j^* \sum_i a_i | e_i \rangle = \sum_j \sum_i b_j^* a_i \langle e_j | e_i \rangle = \sum_j \sum_i b_j^* a_i \delta_{ij} = \sum_j b_j^* a_j \,, \tag{15}$$

which matches up with what we expect: We multiply the corresponding coordinates, conjugating one of them, and then sum the result. Because of the conjugation, the inner product is not commutative, but instead obeys

$$\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^* \,, \tag{16}$$

as we discussed above. Importantly, while we have chosen a particular basis in which to carry out the calculation of the inner product, the result is basis-independent (which is getting to be a familiar refrain).

A useful feature of an orthonormal basis is that we can "pick out" a particular coordinate via the inner product, just as you find the x-component of an ordinary vector by taking its dot product with $\hat{\boldsymbol{x}}$. If we consider the expression for $|\psi\rangle$ in Eq. (13) and take the inner product of both sides with $\langle e_j|$, we obtain

$$\langle e_j | \psi \rangle = a_j \,, \tag{17}$$

since the j^{th} term in the sum gives an inner product of 1 and all the other inner products are zero. Similarly, if Eq. (3) is given in an orthonormal basis we have

$$A_{ij} = \langle e_i | \hat{A} | e_j \rangle \,, \tag{18}$$

and for this reason expressions of this form are often referred to as "matrix elements."

D. Quantum States as Vectors

We will represent the *state of a quantum system* — that is, all the information about the system at given moment in time — as a vector in some (potentially high- or even infinite-dimensional) vectorspace. *Physical quantities* will then be associated with linear operators acting on that vectorspace.

In working with our quantum state vectors, a very useful tool will be the *eigenstates* and *eigenvalues* of a linear operator. An eigenstate (or eigenvector) of an operator $\hat{A}$ is a state $|\lambda\rangle$ such that

$$\hat{A} | \lambda \rangle = c_\lambda | \lambda \rangle \,, \tag{19}$$

where c_λ is a number, which is the corresponding eigenvalue. Such a state is very special: when you act with the operator, you get back to the same state, up to a constant. So the "direction" of the state (or vector) has not changed — it has just been rescaled. (In general, the result of acting with a linear operator would be a linear combination of an entire basis worth of states.)

Eigenstates and eigenvalues will be be crucial in understanding quantum measurements. First, *the result of a measurement is always an eigenvalue of the corresponding operator.* This is what puts the "quantum" in quantum mechanics: The reason there are discrete energy levels in the hydrogen atom is that there only exist eigenstates of the energy operator (which we'll define later) with those eigenvalues; the energy operator has a *discrete* spectrum of possible eigenvalues, and there doesn't exist an eigenstate with an eigenvalue corresponding to the energies in between two of these energy levels. Second, if your system is in a state $|\psi\rangle$, you can predict the *probability* of obtaining a particular result for your measurement from the inner product of $|\psi\rangle$ with the corresponding eigenstate,

$$\text{Probability of measuring } c_\lambda \text{ for system in state } |\psi\rangle = |\langle \lambda | \psi \rangle|^2 \,. \tag{20}$$

(For a continuous physical quantity like x, this becomes a probability *density*.) In general, then, your probability of obtaining the result c_λ depends on how closely your state $|\psi\rangle$ is aligned with the corresponding eigenstate $|\lambda\rangle$. This definition assumes that the state $|\psi\rangle$ is *normalized*,

$$\langle\psi|\psi\rangle = 1 \tag{21}$$

that is, its magnitude (squared) is 1. We have also assumed that the eigenstate $|\lambda\rangle$ is normalized — in particular, it is an element of an orthonormal basis. These assumptions ensure that the total probability of finding *any* outcome for our measurement is 1.

E. Hermitian Operators

Suppose we take a state $|\psi\rangle$, and act on it with a linear operator $\hat{A}$. The result is another state, which we'll call $|\gamma\rangle$,

$$|\gamma\rangle = \hat{A}|\psi\rangle\,. \tag{22}$$

Consider the inner product of this result with another state $\langle\phi|$,

$$\langle\phi|\gamma\rangle = \langle\phi|\hat{A}|\psi\rangle\,. \tag{23}$$

Now we would like to write a similar expression for $\langle\gamma|\phi\rangle$. On the one hand, the inner product is commutative up to conjugation, so for any two states

$$\langle\phi|\gamma\rangle = \langle\gamma|\phi\rangle^*\,. \tag{24}$$

To relate this expression back to ψ, however, we need to introduce a new operator $\hat{A}^\dagger$, which is defined so that $\langle\gamma| = \langle\phi|\hat{A}^\dagger$. Then we have

$$\langle\phi|\hat{A}|\psi\rangle^* = \langle\phi|\gamma\rangle^* = \langle\gamma|\phi\rangle = \langle\psi|\hat{A}^\dagger|\phi\rangle\,. \tag{25}$$

The operator $\hat{A}^\dagger$ is called the *Hermitian conjugate* (or *adjoint*) of $\hat{A}$. We can get some insight into this operator by looking at the case where $|\phi\rangle$ and $|\psi\rangle$ are two elements of an orthonormal basis, which we'll call $|e_i\rangle$ and $|e_j\rangle$. Then we can define the "matrix elements," the entries in the representation of the operator $\hat{A}$ in this basis,

$$\hat{A}_{ij} = \langle e_i|\hat{A}|e_j\rangle\,. \tag{26}$$

From the above, we find that the operator $\hat{A}^\dagger$ is obtained from $\hat{A}$ by switching i and j (that is, taking the transpose of the matrix) and then conjugating,

$$\hat{A}^\dagger_{ij} = \langle e_j|\hat{A}|e_i\rangle^* = \left(\hat{A}_{ji}\right)^*\,. \tag{27}$$

A *Hermitian* (or *self-adjoint*) operator is one for which $\hat{A} = \hat{A}^\dagger$. (In mathematics there is a slight distinction between the terms Hermitian and self-adjoint, but for us they will be equivalent.) Hermitian operators have several extremely useful properties, which we will now discuss (and which you'll prove in the problems):

1. *The eigenvalues of a Hermitian operator are real.*

2. *Eigenstates of a Hermitian operator corresponding to different eigenvalues are orthogonal.*

   Note that if we have several eigenstates of $\hat{A}$ with the *same* eigenvalue, any linear combination of those states is also an eigenstate of $\hat{A}$, also with the same eigenvalue. (This is *not* true for a linear combination of eigenstates with different eigenvalues; in that case a nontrivial linear combination will not be an eigenstate of $\hat{A}$.) So even if we have several independent states with the same eigenvalue that are not orthogonal, we can always choose linear combinations that will be orthogonal. This leads us to the next result:

3. *For any Hermitian operator $\hat{A}$, there exists an orthonormal basis for the space of all quantum states consisting of eigenstates of $\hat{A}$.*

   We have already seen that the eigenstates of $\hat{A}$ can be made orthogonal, and by normalizing them we obtain an orthonormal set. The one thing left to prove is that this orthonormal set spans the space of states, that is, that every state can be written as a linear combination of eigenstates of $\hat{A}$. A rigorous proof hinges on carefully describing the space of allowed states, but for our purposes we will in effect take this set to be defined as the space spanned by our eigenstates, and assume (correctly) that it is sufficient to describe the possible states of the physical system.

F. Unitary Operators

We can think of Hermitian operators as the generalization of real numbers, which makes it appropriate to associate them with physical quantities. We will also be interested in operators where taking the Hermitian conjugate gives the inverse of the original operator, rather than the operator itself. Such operators are called *unitary*, and obey

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}\,, \tag{28}$$

where $\hat{I}$ is the identity operator, which maps every state to itself. While Hermitian operators have real eigenvalues, unitary operators have eigenvalues with magnitude 1, and we can think of unitary operators as generalizations of complex numbers with unit magnitude, $e^{i\theta}$ with θ real. Consistent with this analogy, we will be able to express every unitary operator as $e^{i\hat{H}}$, where $\hat{H}$ is Hermitian (we first will have to define what we mean by the exponential of an operator). As in the case of Hermitian operators, for any unitary operator we will always be able to construct an orthonormal basis consisting of its eigenstates.

G. Commutators

When we have two operators, we can *compose* them, meaning that we act with one and then the other. So

$$\hat{A}\hat{B}|\psi\rangle \tag{29}$$

means we should take $|\psi\rangle$ and then act with $\hat{B}$ first, and then act on the result with $\hat{A}$. Crucially, this is *not* necessarily the same thing as

$$\hat{B}\hat{A}|\psi\rangle\,, \tag{30}$$

where we act with the operators in the other order. In the special case where $\hat{A}\hat{B}|\psi\rangle = \hat{B}\hat{A}|\psi\rangle$ for all states $|\psi\rangle$, we say that $\hat{A}$ and $\hat{B}$ *commute*. Whether operators commute will be of such central interest in quantum mechanics that we define a special symbol representing how much they *fail* to commute, called the commutator:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}\,. \tag{31}$$

Then $\hat{A}$ and $\hat{B}$ commute if and only if $[\hat{A}, \hat{B}] = 0$.

The commutator will show up all over the place, but at this point we'll highlight one more technical result that underlies its many roles. We know that for any Hermitian operator $\hat{A}$, we can construct an orthonormal basis consisting entirely of eigenstates of $\hat{A}$. Now suppose we have two operators, $\hat{A}$ and $\hat{B}$. Can we construct a *single* basis consisting entirely of states that are eigenstates of *both* $\hat{A}$ and $\hat{B}$? (We don't require that the corresponding eigenvalues be equal.) The answer is yes, if and only if $\hat{A}$ and $\hat{B}$ commute. In that case we say that $\hat{A}$ and $\hat{B}$ are *simultaneously diagonalizable*, since in such a basis of eigenstates, both operators would be represented as diagonal matrices. If not, we say that $\hat{A}$ and $\hat{B}$ correspond to *incompatible observables*. That's where things get interesting: if we are considering the operators corresponding to two incompatible observables we have a conflict: In general we can put the system in an eigenstate of one operator or the other, but not both. But eigenstates are the only case where we know the result of a measurement with certainty. So nonvanishing commutators lead to uncertainty, with one example being the famous Heisenberg uncertainty principle.

H. Basis Independence

In quantum mechanics, it turns out to be very useful to work with a quantum state $|\psi\rangle$ in a way that is independent of a particular representation as a wavefunction. The situation is analogous to defining an ordinary 3-dimensional vector $\boldsymbol{v}$ in a way that is independent of a particular coordinate system. The irony is that to actually work with a vector, we always need to choose a particular coordinate system. This means choosing a set of *basis vectors*, like $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$, in terms of which we can represent any vector. However, physical laws will always depend on quantities that are basis-independent, like the magnitude of a vector or the inner product of two vectors, or they will give results that hold equally well for any coordinate system. So it is very valuable to formulate laws of physics in a basis-independent way, whether we are dealing with quantum states or ordinary vectors. (Mathematically, the only difference between these two cases is that the dimension of the space of quantum states — the number of independent coordinates — is *infinite* for most common systems.)

II. THE POSTULATES OF QUANTUM MECHANICS

Let's begin by laying out the underlying assumptions of quantum mechanics, some of which we have already seen. As usual in physics, we can't derive our initial assumptions; we accept them because they lead to correct predictions for what we observe experimentally.

1. The *state of a quantum system* is completely described by a *normalized state vector* $|\psi\rangle$ in some complex Hilbert space. ("Hilbert space" is just the mathematician's term for a possibly infinite-dimensional vectorspace with an inner product. It is also required to be a complete space, which means that the limit of any convergent sequence of points in the Hilbert space is also in the space.) Note that by saying that the state is "completely" described by $|\psi\rangle$, this postulate prevents us from introducing "hidden variables" that one might hope to use to make quantum theory deterministic. In a theory with hidden variables, the outcome of a measurement that looks probabilistic in quantum mechanics is actually governed by additional information beyond the quantum state. It would seem hard to disprove such a claim, but it turns out that under very reasonable assumptions (e.g., only allowing local interactions), it can be shown that all hidden variable theories make predictions that are in conflict with experiment.

2. To every *measurable quantity*, we associate a *Hermitian* operator $\hat{\mathcal{O}}$.

3. If we *measure a particular physical quantity*, the result is always an *eigenvalue* λ of the associated operator $\hat{\mathcal{O}}$. Since the operator is Hermitian, this eigenvalue is always a real number.

4. The *probability of measuring the outcome* λ for the physical quantity associated with the operator $\hat{\mathcal{O}}$ for a system in the state $|\psi\rangle$ is given by the equation $P_\lambda = |\langle\lambda|\psi\rangle|^2$, where $|\lambda\rangle$ is the normalized eigenvector of $\hat{\mathcal{O}}$ with eigenvalue λ. Since $\hat{\mathcal{O}}$ is Hermitian, its eigenvectors form an orthonormal basis, and each probability P_λ is just the magnitude squared of the coordinate of $|\psi\rangle$ corresponding to the basis vector $|\lambda\rangle$. Thus summing the probabilities over all possible outcomes is the same as summing the magnitudes squared of all the coordinates of $|\psi\rangle$ in this basis. Since $|\psi\rangle$ is normalized, the probabilities sum to one (as they should).

5. After performing a measurement whose outcome is λ, the system *suddenly jumps* into the state $|\lambda\rangle$, meaning that a second measurement of the same physical quantity performed directly after the first measurement will always yield the same result.

6. If we don't perform any measurements but simply let time pass, the state of the system *evolves according to the Schrödinger equation*

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{\mathcal{H}}|\psi(t)\rangle\,, \tag{32}$$

   where $\hat{\mathcal{H}}$, called the Hamiltonian, is the Hermitian operator corresponding to the *energy* of the system.

To make sense of these postulates, we'll develop explicit examples to understand how they work and what they mean. To start with, we'll focus on the first four postulates, which concern the description of a state and the results of a measurement of physical quantities. The second two, in contrast, concern dynamics — what happens next. We'll come back to those in due course.

III. EXPLICIT EXAMPLE: 2-DIMENSIONAL VECTORSPACE

To illustrate concretely the concepts we've discussed so far, let's consider a 2-dimensional vectorspace, with basis

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{33}$$

so that a general *column vector* can be written as

$$|\psi\rangle = a|+\rangle + b|-\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \tag{34}$$

where a and b are complex numbers. The inner (dot) product of two vectors is denoted as

$$\langle\psi_2|\psi_1\rangle = \begin{pmatrix} a_2^* & b_2^* \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = a_2^* a_1 + b_2^* b_1 \tag{35}$$

where we have defined the *row vector*

$$\langle\psi| = \begin{pmatrix} a^* & b^* \end{pmatrix} \tag{36}$$

(for mathematical purists: this is an element of the *dual vectorspace*). Note that

$$\langle\psi_2|\psi_1\rangle = \langle\psi_1|\psi_2\rangle^* \tag{37}$$

We can let a linear operator act on our vector

$$|\phi\rangle = \hat{M}|\psi\rangle = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{38}$$

and the corresponding row vector is

$$\langle\phi| = \langle\psi|\hat{M}^\dagger = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix} \tag{39}$$

where we have defined the *adjoint* $\hat{M}^\dagger$, which is the conjugate transpose of the matrix $\hat{M}$.

The inner product of a vector with itself is the norm of that vector squared. In the 2-dimensional case,

$$\langle\psi|\psi\rangle = a^* a + b^* b = ||\psi\rangle|^2 \tag{40}$$

A *quantum state* is represented by a *normalized* vector, so that

$$\langle\psi|\psi\rangle = 1 \tag{41}$$

A *physical quantity* will be represented by a *Hermitian operator*, which is an operator whose adjoint is itself. For the 2 by 2 case,

$$\hat{\mathcal{O}} = \hat{\mathcal{O}}^\dagger = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \tag{42}$$

where A and C are real.

The *eigenvalues and eigenvectors of* $\hat{\mathcal{O}}$ are the states $|\lambda_i\rangle$ such that

$$\hat{\mathcal{O}}|\lambda_i\rangle = \lambda_i|\lambda_i\rangle \tag{43}$$

where λ_i is a number. (The confusing but standard convention is to label the eigenvectors by the corresponding eigenvalues.) We will always choose to normalize the eigenvectors $|\lambda_i\rangle$ so that

$$\langle\lambda_i|\lambda_i\rangle = 1 \tag{44}$$

1. Let $\hat{A}$ and $\hat{B}$ be two linear operators. The *commutator* of $\hat{A}$ and $\hat{B}$ is

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{45}$$

Show that

(a) $[\hat{A}, \hat{A}] = 0$

(b) $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$

(c) $[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]$

2. Consider a linear operator $\hat{A}$, and an eigenvector $|v\rangle$ with eigenvalue λ, so that

$$\hat{A}|v\rangle = \lambda|v\rangle \tag{46}$$

(a) Show $\hat{A}^n|v\rangle = \lambda^n|v\rangle$**.**

(b) Let $f(\hat{A}) = \sum_{n=0}^{\infty} a_n \hat{A}^n$**. Show** $f(\hat{A})|v\rangle = f(\lambda)|v\rangle$**.**

3. Let $|e_1\rangle$ and $|e_2\rangle$ be an orthonormal basis for a vector space. Consider the linear operator $\hat{A}$ defined by

$$\hat{A}|e_1\rangle = |e_2\rangle \qquad \hat{A}|e_2\rangle = -|e_1\rangle \tag{47}$$

(a) Represent $\hat{A}$ **by a matrix in this basis.**

(b) Show that $\hat{A}^4 = \hat{I}$ **where** $\hat{I}$ **is the identity.**

(c) Using (b) show that all the eigenvalues of $\hat{A}$ **are fourth roots of unity (that is, they obey** $\lambda^4 = 1$**).**

(d) Find normalized eigenvectors $|w_1\rangle$ **and** $|w_2\rangle$ **in terms of** $|e_1\rangle$ **and** $|e_2\rangle$**.**

(e) Represent $\hat{A}$ by a matrix acting on the basis $|w_1\rangle$ and $|w_2\rangle$.

4. Suppose that we have $[\hat{A}, \hat{B}] = \hat{I}$. **Show that there is no vector that is simultaneously an eigenvector of $\hat{A}$ and $\hat{B}$.**

   Note: By taking the trace of the above equation, you might conclude that such a situation is impossible. However, it can occur for infinite-dimensional matrices.

5. Consider the matrices

$$\hat{L}_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \hat{L}_y = \frac{1}{i\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \qquad \hat{L}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{48}$$

   (a) Find $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$.

   (b) For each of the operators $\hat{L}^2$, $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, $\hat{L}_x^2$, $\hat{L}_y^2$, and $\hat{L}_z^2$, find a complete set of normalized eigenvectors and their eigenvalues.

   (c) Show explicitly that each set of eigenvectors you found in (b) forms an orthonormal basis.

   (d) Find the commutators $[\hat{L}_x, \hat{L}_y]$, $[\hat{L}_y, \hat{L}_z]$, $[\hat{L}_z, \hat{L}_x]$, $[\hat{L}^2, \hat{L}_x]$, $[\hat{L}^2, \hat{L}_y]$, and $[\hat{L}^2, \hat{L}_z]$.

6. We said that if $\widehat{\mathcal{O}}$ is Hermitian:

   (a) λ_i will always be real.

   (b) Different states $|\lambda_i\rangle$ corresponding to different λ_i will be orthogonal (that is, their inner product is zero).

   (c) The total number of eigenvectors will equal the dimension of the vectorspace (2 in our example), meaning that the eigenvectors form a basis (since they're orthogonal).

   Prove these first two of these three statements. Your proofs should be general, not specific to the 2-dimensional case.

7. The *possible results of a measurement* of the physical quantity are the eigenvalues of $\hat{\mathcal{O}}$.

   For example, consider these three operators:

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{49}$$

   (don't worry about what they actually correspond to physically for now). **For each of the these three operators, find the possible results of a measurement.**

8. The *probability that a given state $|\psi\rangle$ results in the measurement of λ_i* for the physical quantity $\hat{\mathcal{O}}$ is

$$P_i = |\langle \lambda_i | \psi \rangle|^2 \tag{50}$$

Remember that $\langle\lambda_i|\psi\rangle$ is simply the i^{th} coordinate of $|\psi\rangle$ in the basis of eigenvectors $\{|\lambda_i\rangle\}$. As a result, the norm squared of $|\psi\rangle$ is just the sum of all the magnitudes squared of its coordinates, which is just the sum of all the P_i, and so *normalizing* $|\psi\rangle$ *ensures that the probabilities sum to one.*

Start with the state

$$|\psi\rangle = \frac{1}{\sqrt{3}}|+\rangle + i\sqrt{\frac{2}{3}}|-\rangle \tag{51}$$

For each of the operators above, you found the possible results of a measurement. **For a system in the state $|\psi\rangle$, find the probability of each of these results for each operator.**

9. The *expectation value* of the operator $\hat{\mathcal{O}}$ in the state $|\psi\rangle$ is defined to be $\langle\psi|\hat{\mathcal{O}}|\psi\rangle$. In 2 dimensions, we have

$$\langle\psi|\hat{\mathcal{O}}|\psi\rangle = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{52}$$

Show that this quantity represents the *average* result of a measurement of $\hat{\mathcal{O}}$, by which I mean the *average of each possible result of a measurement weighted by the probability of that result.* Hint: Since the eigenvectors of $\hat{\mathcal{O}}$ form a basis, you can write

$$|\psi\rangle = \sum_i c_i|\lambda_i\rangle \tag{53}$$

First find c_i, and then see what happens when you plug this formula into Eq. (52). Again, make your argument apply for arbitrary dimension.

10. **For the state**

$$|\psi\rangle = \frac{1}{\sqrt{3}}|+\rangle + i\sqrt{\frac{2}{3}}|-\rangle \tag{54}$$

compute $\langle\psi|\hat{\sigma}_x|\psi\rangle$, $\langle\psi|\hat{\sigma}_y|\psi\rangle$, and $\langle\psi|\hat{\sigma}_z|\psi\rangle$, where

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{55}$$

Check that this result agrees with the average of the eigenvalues weighted by the corresponding probabilities.

11. **If $\hat{A}$, $\hat{B}$ and $\hat{C}$ are Hermitian, find which of the following are also necessarily Hermitian operators:**

(a) $\hat{A} + \hat{B}$

(b) $\frac{1}{2i}[\hat{A}, \hat{B}] = \frac{1}{2i}(\hat{A}\hat{B} - \hat{B}\hat{A})$

(c) $\hat{A}\hat{B}\hat{C} - \hat{C}\hat{B}\hat{A}$

(d) $\hat{A}^2 + \hat{B}^2 + \hat{C}^2$

(e) $\hat{A} + i\hat{B}$

12. Let $\hat{U}$ be *unitary*, which means that $\hat{U}^\dagger \hat{U} = \hat{1}$.

    (a) **Show that if $\langle\psi|\psi\rangle = 1$, then $\langle\gamma|\gamma\rangle = 1$, where $|\gamma\rangle = \hat{U}|\psi\rangle$.**

    (b) Suppose $|u_i\rangle$ is a complete orthornormal set, so that we have

    $$\langle u_j|u_i\rangle = \delta_{ij} \tag{56}$$

    where the *Kronecker delta function* is one if $i = j$ and zero if $i \neq j$. **Show that $|v_i\rangle = \hat{U}|u_i\rangle$ is also an orthonormal set.**

    (c) Suppose $\hat{U}$ is written out as a matrix in an orthonormal basis. **Show that its columns form an orthonormal set. Show that its rows also constitute an orthonormal set.** Hint: Use the previous result, and also that $\hat{U}$ is unitary if and only if $\hat{U}^\dagger$ is unitary.

13. Consider the matrix $\hat{L}_y$ from above. (You may use your previous results without rederiving them.)

    (a) Find the matrix $\hat{L}'_y$ that represents $\hat{L}_y$ in the basis where it is diagonal.

    (b) Find the matrix $\hat{S}$ such that

    $$\hat{L}'_y = \hat{S}\hat{L}_y\hat{S}^{-1} \tag{57}$$

    (c) Show that $\hat{S}$ is unitary.