Introduction to Vectors and Quantum States

The following continuation of our discussion of the use of vectors to describe quantum states will rely heavily on the following concepts from linear algebra. If you do not yet feel comfortable with these concepts, please brush up on the problematic areas before trying to apply them to the quantum theory.

1. Scalar multiplication of vectors
2. Addition of vectors
3. The inner product of vectors (also called dot product or scalar product)
4. Calculating the length of a vector using inner products
5. Normalizing a vector (making its length 1)
6. Basis vectors
7. Linear combination of vectors
8. Orthogonality and orthonormality
9. Projections of vectors

1 Eigenvectors and Eigenvalues

We mentioned previously that eigenvectors and eigenvalues lie at the heart of the quantum theory. Let's revisit what we have mentioned already but go into a bit more detail.

Consider an $n \times n$ matrix $\mathbf{O}$ that represents a quantum mechanical operator. Because the matrix is $n \times n$, there will be n eigenvectors, each of which comprises n elements. There will also be n eigenvalues.[1] We can express this mathematically with a compact equation:

$$\mathbf{O}\vec{v_i} = \lambda_i \vec{v_i} \tag{1}$$

where the index i simultaneously specifies a particular eigenvector v_i and its matching eigenvalue, λ_i. We said $\mathbf{O}$ is an $n \times n$ matrix so we expect our set

[1] Not all matrices have this property, but all the matrices that represent quantum mechanical operators do.

of eigenvectors to include $\vec{v_1}$, $\vec{v_2}$, $\ldots \vec{v_n}$, and the matching set of eigenvalues to be $\lambda_1, \lambda_2 \ldots \lambda_n$. You should think of each eigenvector and its corresponding eigenvalue as a pair—a matched set.

We said previously that in the quantum theory the states of a system are encoded as vectors. We will now be more specific. It is actually the *orientation* of a vector that encodes a state. That is why it is important to know which mathematical operations leave a vector's orientation unchanged: changing a vector's orientation implies a change in quantum state. Conversely, a mathematical operation that leaves a vector's orientation unchanged corresponds to a quantum mechanical process or event that leaves a quantum state unchanged.

Equation **??** says that when a matrix multiplies one of its eigenvectors the effect is equivalent to multiplying that eigenvector by a scalar (which we call its eigenvalue). But scalar multiplication of a vector only changes the vector's length, not its orientation. (Causing a vector to point in exactly the opposite direction is not considered a change in orientation for our purposes.) We can therefore conclude that when a matrix representing an observable multiplies one of its eigenvectors, it leaves the orientation of the eigenvector unchanged. This, in turn, corresponds to leaving the quantum state unchanged. It is important to remember that *only* if the vector is an eigenvector of the matrix does multiplication by the matrix leave the vector's orientation unchanged. This is a special case. In other cases matrix multiplication changes vector orientation which implies a change in quantum state.

Eigenvectors and eigenvalues are particularly important when we describe making a measurement on a quantum system. Assume for concreteness that our $n \times n$ matrix $\mathbf{O}$ corresponds to an observable property of the system, O. We can now state one of the central ideas in quantum theory. Note that the newly introduced word *eigenstate* simply refers to a state represented by an eigenvector:

> An idealized measurement of property O of a system will yield one of the n eigenvalues of the operator $\mathbf{O}$ as a result. If the eigenvalue λ_i is obtained, then after the measurement the system is left in the corresponding eigenstate, represented by $\vec{v_i}$.

So, if we choose to measure a system's energy, for example, the measurement will yield as a result one of the eigenvalues of the energy operator. After the measurement the system will be left in the energy eigenstate corresponding to that eigenvalue. If we choose to measure some other property, such as position, we will obtain a position eigenvalue and force the system into a position eigenstate. As we will see later, position eigenstates cannot also be energy eigenstates so if we first measure energy and then measure position we can be sure that we are forcing the system to change states as a result of our prodding.

2 Applications

2.1 Orbitals and Energy Levels

You are already familiar with some of these concepts, but may not have encountered them expressed in the language of linear algebra. For example, the energies of the levels available to an electron in a hydrogen atom are the *eigenvalues* of the energy operator for a hydrogen atom. The orbitals of the hydrogen atom are the states with well defined energy. They are the *energy eigenstates*—the states represented by eigenvectors of the hydrogen atom energy operator.

Since there are an infinite number of energy levels and orbitals, the energy operator for the hydrogen atom must be represented by an $\infty \times \infty$ matrix. Fortunately we can usually truncate this matrix when tackling chemical problems, and focus on a small subset of the energy levels and orbitals (energy eigenvalues and eigenstates).

2.2 Electron Spin

We are now ready to apply all this to a (superficially) simple system. Every chemistry student has encountered the fact that electrons are said to have two spin states with well-defined energy, "spin up" and "spin down." In the simplest level of description we think of these two states as corresponding to the particle's magnetic dipole being aligned with or against an external magnetic field, respectively.[2] Because there are only two states with well-defined energy we can immediately draw the following conclusions:

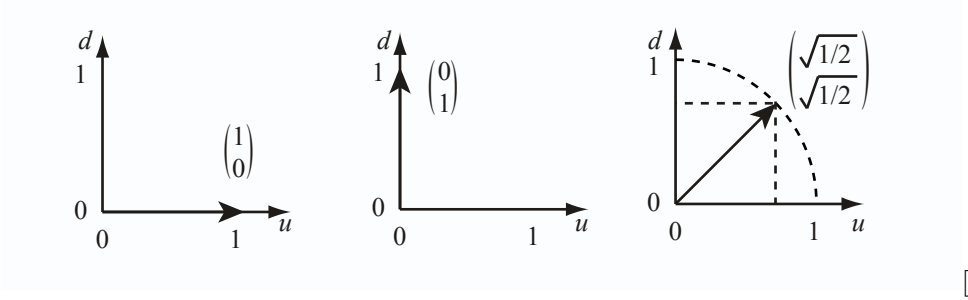
- Energy in this system is represented by a 2×2 matrix
- There are two energy eigenvectors
- Each eigenvector has two elements
- There are two energy eigenvalues.

We can now rewrite our eigenvector equation as it applies to this situation. First we adopt the following mathematical notation:

$\mathbf{H}$	total energy operator, or hamiltonian operator
$\vec{\alpha}$	vector representing a spin up electron
λ_α	energy of an electron in the spin up state
$\vec{\beta}$	vector representing a spin down electron
λ_β	energy of an electron in the spin down state

The eigenvector-eigenvalue pairs are regrouped to emphasize they should be thought of together. Rewriting equation **??** separately for each of the two eigen-

[2] Later we will refine this description of the orientation of a magnetic moment with respect to particular axes.



[b]

Figure 1: Three of the many possible spin states depicted as arrows. The horizontal axis, labeled u, corresponds to spin up, the vertical axis, labeled d, corresponds to spin down. The first state can be represented by $\vec{\alpha} = (1, 0)$, the second by $\vec{\beta} = (0, 1)$, and the third by $\vec{\psi} = (1/\sqrt{2}, 1/\sqrt{2})$

vectors and eigenvalues yields

$$\mathbf{H}\vec{\alpha} = \lambda_\alpha \vec{\alpha} \tag{2}$$

$$\mathbf{H}\vec{\beta} = \lambda_\beta \vec{\beta} \tag{3}$$

From the previous discussion we know that equation **??** implies that if we place an electron in the spin state $\vec{\alpha}$ and measure its energy, we will definitely obtain the energy value λ_α. Equation **??** implies that if the electron is in the state $\vec{\beta}$ and we measure its energy we can be 100% confident that we will obtain the energy value λ_β.

2.2.1 Graphical Representation

To highlight some key features of $\vec{\alpha}$ and $\vec{\beta}$ we can use the common "arrow" representation of vectors. Since our vectors each have two elements, we must use a two-dimensional space to draw them.

In this example we will define our coordinates so that our $\vec{\alpha}$ vector lies along the horizontal axis (labeled u for up) and $\vec{\beta}$ lies along the vertical axis (labeled d for down). We could have chosen other coordinates, as we will see later, but these are convenient for now.

If we normalize our eigenvectors then we can be confident that they have length 1. The vector $\vec{\alpha}$ should be as shown in the left diagram in figure **??**. Notice it has length 1 and that in this ud coordinate system it can be written (1, 0). This states that $\vec{\alpha}$ has projections of 1 on u and 0 on d. Similarly, $\vec{\beta}$ is as shown in the middle diagram.

The point you should now pause to digest is that $\vec{\alpha}$ and $\vec{\beta}$ are the only two energy eigenvectors for this system but they are emphatically *not* the only two

possible states of the system.

Consider another vector in the same *ud* space. Let's call it $\vec{\psi}$. It is shown in the third diagram in figure **??**. Since $\vec{\psi}$ exists in the same space as $\vec{\alpha}$ and $\vec{\beta}$, and since $\{\vec{\alpha}, \vec{\beta}\}$ form a basis (which the figure shows to be orthonormal), we must be able to represent $\vec{\psi}$ as a linear combination of $\vec{\alpha}$ and $\vec{\beta}$. The figure shows $\vec{\psi}$ lying midway between $\vec{\alpha}$ and $\vec{\beta}$ so it must have equal amounts of $\vec{\alpha}$ and $\vec{\beta}$ "character." This is captured by the fact that when we write $\vec{\psi}$ as a column of numbers in the $\vec{\alpha}$, $\vec{\beta}$ basis, the two elements of $\vec{\psi}$ have the same absolute value. The graph also shows that $\vec{\psi}$ has length 1. Applying what we learned about calculating the lengths of vectors we conclude that $\vec{\psi}$ can be written in this *ud* space as $(\sqrt{1/2}, \sqrt{1/2})$. The next section reminds you how to calculate vector lengths in case you need a reminder.

To reinforce the very important connections between the arrow graphics, the expression of vectors as columns of numbers, and the use of linear combinations of basis vectors, take the time to verify that the following expressions are both internally consistent and also consistent with the graphs in figure **??**. Don't move on to the next section until you see clearly how the elements comprised by the vectors, the expansion coefficients, and the projections in the graphs all say the same things!

$$\vec{\alpha} \quad = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot \vec{\alpha} + 0 \cdot \vec{\beta} \quad = 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4}$$

$$\vec{\beta} \quad = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \cdot \vec{\alpha} + 1 \cdot \vec{\beta} \quad = 0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5}$$

$$\vec{\psi} \quad = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \tfrac{1}{\sqrt{2}} \cdot \vec{\alpha} + \tfrac{1}{\sqrt{2}} \cdot \vec{\beta} \quad = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6}$$

3 Lengths of Vectors and the Inner Product

We emphasized that it is the orientation of a vector that encodes a quantum state. The vectors $\vec{\alpha}$, $2\vec{\alpha}$, and $-5\vec{\alpha}$ all represent the same state because the vector orientations are the same; only the lengths are different. Nonetheless, we will have to pay attention to vector lengths in some situations. We have already discussed the use of the Pythagorean theorem as one way to measure the length of a vector, and then introduced the inner product as a more compact and easily generalizable approach. We revisit them here briefly so the key ideas are fresh.

To use the Pythagorean formula to measure vector length consider the arrow representing a vector to be the hypotenuse of a right triangle. If we know the length of the arrow's projection onto each of the perpendicular axes defining the coordinate system we can easily calculate the vector's length. The length of a vector $\vec{v} = (u, d)$, written $\| \vec{v} \|$, is given by $\| \vec{v} \| = \sqrt{u^2 + d^2}$. By rewriting the squared terms as products we can emphasize an important feature of the length

calculation. Both forms are shown below.

$$\| \vec{v} \| = \sqrt{u^2 + d^2} = \sqrt{uu + dd} \tag{7}$$

The last form of the expression, in which uu and dd are used rather than u^2 and d^2, can be read as a manifestation of our mantra "multiply corresponding elements and add up," which you should recognize as the multiplication of vectors. This serves to remind us that a more compact notation is available:

$$\| \vec{v} \| = \sqrt{\langle \vec{v}, \vec{v} \rangle} \tag{8}$$

where the angle bracket notation indicates an "inner product" or "scalar product" of vectors. To compute an inner product we must perform three operations.[3] First, transpose the vector written on the left so that it becomes a row vector. Then take its complex conjugate, which simply means replace $i = \sqrt{-1}$ with $-i = -\sqrt{-1}$ everywhere i appears. Here we're only using real numbers so complex conjugation leaves the left vector unchanged. Finally, perform the usual vector multiplication. Writing out the steps we obtain

$$\| \vec{v} \| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{\begin{pmatrix} u^\star & d^\star \end{pmatrix} \begin{pmatrix} u \\ d \end{pmatrix}} = \sqrt{u^\star u + d^\star d} \tag{9}$$

where the star superscript implies complex conjugation. Notice that we end up with the same result we obtained using Pythagoras's approach. When we have vectors with more than two elements, the advantage of using the inner product will become obvious. For now, to be sure you're keeping track of everything, try applying the inner product method to determine the lengths of the three vectors shown in figure **??**. If all goes well you should find that the lengths of all three vectors are 1.

3.0.1 Normalization

We can make the length of any vector 1 by dividing the vector by its own length. This process is called *normalization* and we will perform it often. Given a vector $\vec{v}$ with length $\| \vec{v} \|$ we can create a new vector $\vec{w}$ with the same orientation as $\vec{v}$ but with length 1:

$$\vec{w} = \frac{\vec{v}}{\| \vec{v} \|}$$

Since the new vector has the same orientation as the original one, both vectors represent the same state. However, the new normalized vector has the advantage that it can be used directly to calculate the probabilities of measurement results—our next topic.

[3] Notice that in this case we're taking the inner product of the vector $\vec{v}$ with itself, but we can also take the inner product of two different vectors, as long as they have the same number of elements.

4 Probabilities of Measurement Outcomes

We know from previous discussions that when observable O is measured, the result obtained will always be one of the eigenvalues of the matrix $\mathbf{O}$. If we represent the state of the system using a normalized vector, we can say not only what the possible measurement results are, but also what the *probability* of obtaining each result is.

Consider a spin 1/2 particle in the state $\vec{\psi}$. The first step when calculating the probability of obtaining λ_α as an energy measurement is to express $\vec{\psi}$ as a linear combination of energy eigenvectors. Recall that we established earlier that those energy eigenvectors ($\vec{\alpha}$ and $\vec{\beta}$) are orthonormal so we are all set to use them. Our linear combination can be written

$$\vec{\psi} = c_1\vec{\alpha} + c_2\vec{\beta} \tag{10}$$

where c_1 and c_2 are the expansion coefficients determining how much the basis vectors contribute to $\vec{\psi}$.

To find the values of c_1 and c_2 we calculate inner products of each basis vector with the vector describing the state of our particle ($\vec{\psi}$).

$$c_1 = \langle\vec{\alpha}, \vec{\psi}\rangle \tag{11}$$

$$c_2 = \langle\vec{\beta}, \vec{\psi}\rangle \tag{12}$$

The probability of obtaining λ_α as a measurement result is given by squaring c_1, but being careful to take into account that it might be complex valued rather than just real. We therefore say that the probability of obtaining λ_α as an energy measurement result is given by

$$\text{Prob}_\alpha = \left|\langle\vec{\alpha}, \vec{\psi}\rangle\right|^2 \tag{13}$$

$$= |c_1|^2 \tag{14}$$

$$= c_1^* c_1 \tag{15}$$

Similarly, to calculate the probability that an energy measurement will yield the result λ_β we determine the value of

$$\text{Prob}_\beta = \left|\langle\vec{\beta}, \vec{\psi}\rangle\right|^2 \tag{16}$$

$$= |c_2|^2 \tag{17}$$

$$= c_2^* c_2 \tag{18}$$

These probability calculations work if and only if we use normalized vectors. Doing so ensures that when we sum the probabilities of all the possible measurement outcomes we get 1. That should make sense: you will definitely get some result when you measure an observable so the sum of the probabilities of all the possibilities should be 1. If we accidentally use *un*normalized basis vectors then we will likely find that the sum of all the probabilities is something other than 1, which is nonsensical.

4.1 An Exercise

To demonstrate for yourself that the calculations work as advertised, consider three possible initial spin states, $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\psi}$, as previously defined. For each initial state calculate the probability that a measurement would yield λ_α and the probability that λ_β would be obtained. Verify that the sum of the probabilities of obtaining λ_α and λ_β is 1. Also verify that if the initial state of the system is an eigenstate of the operator, the outcome of the measurement can be predicted with certainty (probability = 1). Finally, show that if the initial state is not an eigenstate of the operator, the outcome of the measurement can not be predicted with certainty.

4.2 More Dimensions

Usually we will work with observables with more than two dimensions. Let's see how we can extend what we have done with spin to linear position. How would we calculate the probability of finding a particle at specific points along a line? We will need to write the particle's state as a linear combination of the eigenvectors of the pertinent operator, in this case position. Let's use $\mathbf{X}$ to represent the matrix for the observable "position along the x axis." We can then write the eigenvector-eigenvalue equation:

$$\mathbf{X}\vec{x}_i = \lambda_i \vec{x}_i \tag{19}$$

The set of position eigenvectors $\{\vec{x}_i\}$ represents the states with well-defined x-position. The eigenvalues $\{\lambda_i\}$ form the set of possible locations where the particle could be found.

Now consider some state of interest $\vec{\psi}$. We want to know the possibility of finding a particle in the state $\vec{\psi}$ at each of the possible x-positions, x_i. We therefore write $\vec{\psi}$ as a linear combination of position eigenvectors:

$$\vec{\psi} = c_1\vec{x}_1 + c_2\vec{x}_2 + c_3\vec{x}_3 \ldots \tag{20}$$

$$= \sum_i c_i \vec{x}_i \tag{21}$$

The probability of obtaining the ith measurement result is given by $|c_i|^2$, which we remember requires us to multiply c_i^* by c_i in case c_i is complex. In other words, if we want to know the probability of finding the particle in state $\vec{\psi}$ at the position x_3, we calculate $|c_3|^2$ which we have seen is just $|\langle \vec{x}_3, \vec{\psi}\rangle|^2$. After performing the position measurement and obtaining the result x_3 the system will no longer be in the state $\vec{\psi}$. It will instead be left in the eigenvector corresponding to the eigenvalue x_3: $\vec{x}_3$.

4.3 A Probabilistic Theory

Based on the examples just explored it should be clear that the quantum theory only allows us to predict the outcome of an individual measurement if the system

is in an eigenstate of the operator corresponding to the measured property. In those special cases we know which of the eigenvalues will be obtained. In the other more common cases, we can not predict the outcome of an individual quantum measurement. We know what the set of possible measurement results is (the eigenvalues), and we can predict the probability of obtaining each one (by applying equations with the form of equation **??**). Nonetheless, the fact remains that except in the special case of the system starting out in an eigenstate of the property of interest, we can not predict with certainty which of the possible measurement results we will obtain.

4.4 Expectation Value

Fortunately, all is not lost. Since we know the probability of obtaining each measurement result, we can also predict the *average* result of many measurements, as long as we're careful to return the system to the same initial state before each measurement. Perhaps more importantly, if we have many identical copies of a system—many hydrogen atoms in the same state, for example—and we make a measurement on the whole collection, we can predict the measurement result that will be obtained from the entire collection. Again, it's the average of the individual results.

The average of many measurements of the observable O is called the *expectation value* of O and it is symbolized $\langle \mathbf{O} \rangle$. In the case of electron spin we write for the expectation value of energy:

$$\langle \mathbf{H} \rangle = \text{Prob}_\alpha \lambda_\alpha + \text{Prob}_\beta \lambda_\beta \tag{22}$$

$$\langle \mathbf{H} \rangle = \left| \langle \vec{\psi}, \vec{\alpha} \rangle \right|^2 \lambda_\alpha + \left| \langle \vec{\psi}, \vec{\beta} \rangle \right|^2 \lambda_\beta \tag{23}$$

where equation **??** makes use of equations **??** and **??**.

We can write a more concise and general expression in place of equation **??**. Here $\vec{\varphi_j}$ is the jth basis vector (the jth eigenvector of the matrix $\mathbf{O}$), $\vec{\psi}$ is the vector describing the system's state, and λ_j is the jth eigenvalue of $\mathbf{O}$.

$$\langle \mathbf{O} \rangle = \sum_j |\langle \vec{\psi}, \vec{\varphi_j} \rangle|^2 \lambda_j \tag{24}$$

$$\langle \mathbf{O} \rangle = \sum_j |c_j|^2 \lambda_j \tag{25}$$

In equation **??** we use the expansion coefficients c_j to replace the result of the inner products in equation **??**. Since $|c_j|^2$ gives us the probability that the jth eigenvalue will obtained as a measurement result, and the λ_j's are the possible measurement results, $\sum_j |c_j|^2 \lambda_j$ predicts the average of many measurements (assuming the system is always in the state ψ before each measurement is made).

You might wonder why it's of interest to compute average measurement results. In fact, it is extremely important and useful. When we cannot predict with certainty what the result of an individual measurement will be being able to predict averages of many such measurements becomes the next best thing.