Chapter 11

Noncommuting Operators and Uncertainty

11.1 Eigenstates and Commuting Operators

You are probably used to the idea that multiplication is *commutative*. That is, if you have a product ab where a and b are scalars, you can write the multiplication in either order (ab or ba), and the product is exactly the same. This is *not* necessarily the case for matrix multiplication! If A and B are matrices, then $AB \neq BA$ in general. Sometimes it will be true, but not always. Because we can use matrices to represent operators in quantum mechanics, this means that operators don't commute in general. That is, for example, $\hat{S}_x\hat{S}_y |\psi\rangle \neq \hat{S}_y\hat{S}_x |\psi\rangle$.

Sometimes, however, operators do commute. Suppose that you have two observables A and B with corresponding operators $\hat{A}$ and $\hat{B}$. Suppose also that you have a state $|\phi\rangle$ that is a definite state for *both* A and B. That means that in our mathematical formalism, $|\phi\rangle$ must be an eigenvector for both $\hat{A}$ and $\hat{B}$:

$$\hat{A} |\phi\rangle = a |\phi\rangle$$
$$\hat{B} |\phi\rangle = b |\phi\rangle$$

Here, a and b are the eigenvalues for $\hat{A}$ and $\hat{B}$ respectively. In other words, $|\phi\rangle$ has a definite value of observable A, and a is that value; likewise, it has a definite value of observable B, and b is that value.

Let us now consider the application of both of these operators to this state $|\phi\rangle$:

$$\begin{aligned}\hat{A}\,\hat{B} |\phi\rangle &= \hat{A}\,b |\phi\rangle \\ &= b\,\hat{A} |\phi\rangle \\ &= ba |\phi\rangle \\ &= ab |\phi\rangle\end{aligned}$$

where in the last step we've used the fact that a and b are real numbers, so the product of the two of them does in fact commute. Let's now try this in the other order:

$$\begin{aligned} \hat{B}\,\hat{A}\,|\phi\rangle &= \hat{B}\,a\,|\phi\rangle \\ &= a\,\hat{B}\,|\phi\rangle \\ &= ab\,|\phi\rangle \end{aligned}$$

Here, we can see that in fact the operators $\hat{A}$ and $\hat{B}$ *do* commute if they are operating on a state that is an eigenstate for both operators.

Remember that in the case of spin, we argued that $|+z\rangle$ and $|-z\rangle$ form a complete basis set of vectors; that is, any spin state $|\psi\rangle$ can be written as a sum of scalar constants times those two vectors. In general, a complete set of eigenvectors for a given operator do form a basis set that can be used to construct any vector that is part of the overall scheme that that operator is part of. (For instance, the projection of spin along all three axes are part of the same scheme, as they are all the same kinds of states— that is, spin angular momentum states.) Therefore, we can write any state $|\psi\rangle$ as a sum of constants times the eigenvectors for that operator. If $\hat{A}$ and $\hat{B}$ are two operators that share the same eigenvectors, then $\hat{A}\hat{B}\,|\psi\rangle = \hat{B}\hat{A}\,|\psi\rangle$. That is, the operation of these two operators on *any* state commutes. For that reason, we generally just say that the operators commute.

11.2   Non-Commuting Operators

In the previous section, we saw that if a particle can be in a definite state for two observables, then the two operators associated with those observables will commute. The converse is therefore also true; if two operators do *not* commute, then it is *not possible* for a quantum state to have a definite value of the corresponding two observables at the same time.

We've already seen examples of this. A particle can't have a definite x spin and a definite y spin at the same time. If our theory is to be useful, then we would hope that $\hat{S}_x$ and $\hat{S}_y$ would not commute when they operate on a general normalized state

$|\psi\rangle$. Let's try it first in one order:

$$\hat{S}_x \hat{S}_y \,|\psi\rangle \;=\; \frac{\hbar^2}{4} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

$$= \;\frac{\hbar^2}{4} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -i\psi_2 \\ i\psi_1 \end{bmatrix}$$

$$= \;\frac{\hbar^2}{4} \begin{bmatrix} i\psi_1 \\ -i\psi_2 \end{bmatrix}$$

$$= \;i\frac{\hbar^2}{4} \begin{bmatrix} \psi_1 \\ -\psi_2 \end{bmatrix}$$

Now let's try it in the other order:

$$\hat{S}_y \hat{S}_x \,|\psi\rangle \;=\; \frac{\hbar^2}{4} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

$$= \;\frac{\hbar^2}{4} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \psi_2 \\ \psi_1 \end{bmatrix}$$

$$= \;\frac{\hbar^2}{4} \begin{bmatrix} -i\psi_1 \\ i\psi_2 \end{bmatrix}$$

$$= \;-i\frac{\hbar^2}{4} \begin{bmatrix} \psi_1 \\ -\psi_2 \end{bmatrix}$$

Clearly the two are not the same; one is the negative of the other. Therefore, $\hat{S}_x$ and $\hat{S}_y$ do *not* commute when operating on a general state ψ, as expected.

It is interesting to note the effect of $\hat{S}_z$ on this same general state:

$$\hat{S}_z \,|\psi\rangle \;=\; \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

$$= \;\frac{\hbar}{2} \begin{bmatrix} \psi_1 \\ -\psi_2 \end{bmatrix}$$

Notice that except for the constant out front, the vector produced by $\hat{S}_z$ on this state is the same as the vector produced by $\hat{S}_x \hat{S}_y$ and $\hat{S}_y \hat{S}_x$. In fact, we can put the two together:

$$(\hat{S}_x \hat{S}_y \,-\, \hat{S}_y \hat{S}_x) \,|\psi\rangle = i\frac{\hbar^2}{2} \,|\psi\rangle$$

$$[\hat{S}_x , \, \hat{S}_y] \,|\psi\rangle \;=\; i\hbar \, \hat{S}_z \,|\psi\rangle$$

The term in brackets, $[\hat{S}_x, \hat{S}_y]$ is called the *commutator* of $\hat{S}_x$ and $\hat{S}_y$. It's defined by the term in parentheses above it: $(\hat{S}_x \hat{S}_y - \hat{S}_y \hat{S}_x)$. It works out for the commutators of all three spin angular momentum operators that:

$$[\hat{S}_x , \, \hat{S}_y] \;=\; i\hbar \, \hat{S}_z$$

$$[\hat{S}_y, \hat{S}_z] = i\hbar\,\hat{S}_x$$
$$[\hat{S}_z, \hat{S}_x] = i\hbar\,\hat{S}_y$$

11.3   Quantifying Uncertainty

If a system is in an indeterminate state for a given observable, it means that we can't know exactly what we're going to measure if we do in fact make a measurement of that observable. We've seen that we can calculate the average of all the measurements we might make, suitably weighted by their probabilities; that's what is called the "expectation value" in quantum mechanics. And, we have seen how we can calculate the amplitude, and from that the probability, that we'll get any given possible measurement. With the spin-1/2 system we've been talking about, there are only two possible values that we might measure for the spin along any given axis: $+\hbar/2$ and $-\hbar/2$. As such, it's convenient just to list the probability of each. If you talk about other systems where there are a lot of possible measurements (including continuous systems such as the position of a particle), it becomes impractical to list the probabilities of each state. It would be nice to have some other way of quantifying our uncertainty.

11.3.1   Mean and Variance

Suppose you have a set of values a_j. By saying that this is a set, we mean that we have several values a_1, a_2, a_3, and so forth. The notation a_j, in this context, means that j can be replaced by any integer between 1 and the total number of values that you have in order to refer to that specific value. Suppose that we have N total values. The average of all of our values can be written as:

$$\langle a \rangle = \frac{1}{N} \sum_j a_j$$

The letter Σ is the capital Greek letter "sigma". This notation means that you sum together all of the values of a_j that you have. For instance, suppose you had just four values, a_1, a_2, a_3, and a_4, then:

$$\sum_j a_j = a_1 + a_2 + a_3 + a_4$$

Therefore, the mean (or average) value of a in this context is:

$$\langle a \rangle = \frac{1}{N} \sum_j a_j = \frac{1}{N}(a_1 + a_2 + a_3 + a_4)$$

To quantify the uncertainty on a set of values, we want to say something about how far, on average, a given value is from the mean of all the values. Thus, it's tempting to try to define the uncertainty as follows:

$$\frac{1}{N} \sum_j (a_j - \langle a \rangle)$$

Remember that addition is commutative. Realizing that the $\sum$ symbol just indicates a sum, i.e. a whole lot of addition, we can rewrite this as:

$$\frac{1}{N} \left(\sum_j a_j - \sum_j \langle a \rangle \right)$$

The second term in the subtraction is a sum over j of the average value. The average value doesn't depend on which a_j we're talking about; it's a constant, it's the same for all of them. Therefore, the sum of that number N times is just going to be equal to $N \langle a \rangle$. Making this substitution and distributing the $1/N$ into the parentheses:

$$\frac{1}{N} \sum_j a_j - \frac{1}{N} N \langle a \rangle$$

But we recognize the first term in this subtraction as just $\langle a \rangle$. So, the total result of this is zero. Clearly, this is not a good expression for the uncertainty in a. If you think about it, the average deviation of a_j from $\langle a \rangle$ *ought* to be zero. If $\langle a \rangle$ is the average value of a, then a_j should be below $\langle a \rangle$ about as often as it is above, so your sum will have a mix of positive and negative terms. The very definition of the average insures that this sum will be zero.

Instead, we shall define the *variance* as:

$$\Delta a^2 = \frac{1}{N} \sum_j (a_j - \langle a \rangle)^2$$

Here, we're using Δa to indicate the uncertainty in a. The variance is defined as the uncertainty squared.[1] The advantage of this expression is that because we're squaring the difference between each value a_j and the average value, we're always going to be summing together positive terms; there will be no negative terms to cancel out the positive terms. Therefore, this should be a reasonable estimate of how far, typically, the measurements a_j are from their average.

---

[1]If you know statistics, you may recognizing this as being very similar to how variance is defined there— only in statistics, we divide by $N-1$ rather than by N. The difference becomes unimportant as N gets large.

We can unpack this sum a bit, first by multiplying out the squared polynomial:

$$\Delta^2 = \frac{1}{N} \sum_j (a_j^2 - 2\langle a \rangle\, a_j + \langle a \rangle^2)$$

In order to clean this expression up, inside the parentheses both add and subtract $\langle a \rangle^2$:

$$\Delta a^2 = \frac{1}{N} \sum_j (a_j^2 - 2\langle a \rangle\, a_j + 2\langle a \rangle^2 - \langle a \rangle^2)$$

$$= \frac{1}{N} \sum_j (a_j^2 - \langle a \rangle^2 + 2\langle a \rangle\, (\langle a \rangle - a_j))$$

$$= \frac{1}{N} \sum_j a_j^2 - \frac{1}{N} \sum_j \langle a \rangle^2 + \frac{1}{N} 2\langle a \rangle \sum_j (\langle a \rangle - a_j)$$

Notice that the last term is going to be zero, as it includes the average difference between the mean and each observation. The second term is just going to be $\langle a \rangle^2$, because once again $\langle a \rangle$ is the same for all terms of the sum; the sum will yield $N \langle a \rangle^2$, canceling the N in the denominator. So, we have:

$$\Delta a^2 = \langle a^2 \rangle - \langle a \rangle^2$$

11.3.2   Uncertainty in Quantum Mechanics

In order to bring this into quantum mechanics, we already know how to calculate the average $\langle a \rangle$, which we call the "expectation value". If the state of the system is $|\psi\rangle$ and the operator corresponding to the observable a is $\hat{A}$, then

$$\langle a \rangle = \left\langle \psi \,\middle|\, \hat{A} \,\middle|\, \psi \right\rangle$$

Similarly, now that we recognize that we can interpret $\hat{A}^2$ as just applying the operator $\hat{A}$ twice, we can calculate $\langle a^2 \rangle$:

$$\langle a^2 \rangle = \left\langle \psi \,\middle|\, \hat{A}^2 \,\middle|\, \psi \right\rangle$$

For example, let's consider the state $|\psi\rangle = |+z\rangle$ and the observable spin-z. We expect the uncertainty here to be zero, because we know exactly what we'll get if we

measure spin-z. Let's see if it works out that way:

$$
\begin{aligned}
\langle s_z \rangle &= \left\langle \psi \left| \hat{S}_z \right| \psi \right\rangle \\
&= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar}{2}
\end{aligned}
$$

As expected, the expectation value for spin-z is $+\hbar/2$. For the other part:

$$
\begin{aligned}
\langle s_z{}^2 \rangle &= \left\langle +z \left| \hat{S}_z \hat{S}_z \right| -z \right\rangle \\
&= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar^2}{4}
\end{aligned}
$$

If we take the difference $\langle s_z{}^2 \rangle - \langle s_z \rangle^2$, we get $\hbar^2/4 - \hbar^2/4 = 0$, as expected.

What if we want to know the uncertainty on S_x for this state?

$$
\begin{aligned}
\langle s_x \rangle &= \left\langle +z \left| \hat{S}_x \right| +z \right\rangle \\
&= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= 0
\end{aligned}
$$

If the system is in the state $|+z\rangle$, we know that we have a 50% chance each for finding spin-x to be $+\hbar/2$ or $-\hbar/2$. Thus, it's no surprise that the average value of spin-x is

zero, even though zero isn't a value we might measure. To figure out the variance:

$$
\begin{aligned}
\left\langle s_x{}^2 \right\rangle &= \left\langle +z \left| \hat{S}_x \hat{S}_x \right| +z \right\rangle \\
&= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \frac{\hbar^2}{4} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \frac{\hbar^2}{4}
\end{aligned}
$$

Thus, in this case, the formal uncertainty Δs_x on the x-spin is $\hbar/2$.

11.4   The Heisenberg Uncertainty Principle

As previously stated, quantifying the uncertainty on a given observable for a given quantum state is more interesting when the observable we're talking about has a large number (or even a continuum) of different values it might take on. If you consider two different observables whose operators do not commute, then a system cannot be in a definite state for both of those observables at the same time. The Heisenberg Uncertainty Principle takes this observation, makes it stronger, and quantifies it.

Consider a quantum particle that can move along one direction. Its position is then x, and its momentum along that direction is p_x. The Heisenberg Uncertainty Principle states that:

$$
\Delta x \, \Delta p_x \;\geq\; \frac{\hbar}{2}
$$

We've gotten used to thinking of $\hbar$ as an angular momentum unit, because that's where it's shown up before. However, here, it's not really an angular momentum, though it still does of course have the same units (position times position over time). Instead, it represents the fundamental limit in quantum mechanics on how well you can know two different observables, position and momentum. If you know one of them *perfectly*, e.g. $\Delta x = 0$, then the uncertainty in the other one must be *infinite*. Although in more advanced quantum mechanics we use such states as they're a good approximation for a lot of things, they're not really physical. In reality, most quantum systems have a small amount of uncertainty in both position and momentum. That is, a particle doesn't have a definite position or a definite momentum, but the range of positions for which it has an appreciable amplitude is confined to a small space,

and the range of momenta for which it has an appreciable amplitude is confined to a small range.

As a concrete example, let's consider an electron. For a non-relativistic electron, its momentum is just $p = mv$, where m is its mass and v is its speed. Therefore, $\Delta p = m\Delta v$, as the mass is well known and there is no uncertainty in it. What is a good uncertainty in speed to consider? For practical purposes, let's suppose that we're doing an experiment with an electron that requires it to be localized for 1 second. We don't want the uncertainty in the speed of the electron to cause our uncertainty in the position after one second to be greater than the uncertainty in the position was in the first place. So, we shall choose $\Delta v = \Delta x/t$, where we'll put in $t = 1$ second. If we then put this into the uncertainty principle

$$\Delta x \, \Delta p \; \geq \; \frac{\hbar}{2}$$

$$\Delta x \, m_e \, \frac{\Delta x}{t} \; \geq \; \frac{\hbar}{2}$$

$$\Delta x \; \geq \; \sqrt{\frac{\hbar \, t}{2 \, m_e}}$$

If you put in the numbers, you find that the uncertainty on the position of this electron is 0.01 m, or one centimeter. For an electron, that's a lot! (One could argue about whether or not 1 second is a reasonable timescale. When we get to talking about atoms, we'll think more carefully about what a reasonable timescale is.)

Notice, however, that the uncertainty in the position goes down as the mass goes up. Imagine that you stood still your entire life. If you want to balance the uncertainty in your starting position with the uncertainty in your position resulting from the uncertainty in your velocity over your entire life, then you'd put in your age for t. Let us assume, optimistically, that you will live 100 years (3×10^9 seconds), and that your mass is 80 kg. If you put those numbers in to the equation above, you find out that the uncertainty on your position is 4×10^{-14} m. In other words, even though quantum uncertainty can be pretty important for an electron, on everyday scales for macroscopic objects the effect of quantum uncertainty is utterly negligible.

11.4.1   An Alternate Formulation

An alternate way to formulate Heisenberg's Uncertainty principle is:

$$\Delta E \, \Delta t \; \geq \; \frac{\hbar}{2}$$

The interpretation of this is a little less clear than in the case of position and momentum. Uncertainty in energy seems obvious enough; it's the square root of the

variance of all the values of energy that might be measured for a particle in a given quantum state. But what is "uncertainty on time"? Rather than interpreting this as an uncertainty, we shall interpret it as a time interval. In a sense, that's the same thing; an interval of time is qualitatively similar to an uncertainty on what time it is.

What this means, then, is that the uncertainty in the energy of a quantum state is related to how long that state hangs around. If a system is in an energy eigenstate, then it has a definite energy and $\Delta E = 0$. Such a state *must* be stable then, for Δt has to be infinite. In other words, in the absence of any interactions, a particle in an energy eigenstate will stay, forever, in an energy eigenstate.

For small time intervals, however, there will be a finite uncertainty in the energy of a system. One thing that this means is that it becomes possible to violate the conservation of energy, so long as you do it so fast that nobody can catch you at it! Among other things, this leads to the possibility of quantum tunneling– that is, if a particle is up against a potential barrier it doesn't have enough energy to penetrate, there *is* some finite probability that the particle may be located *inside the barrier*. And, the particle may be able to cross the barrier, even though classically it could not.

Later, when we talk about atoms, states other than the ground state (i.e. lowest-energy state) of the atom aren't going to be perfectly stable. Over time, they will decay to the ground state, with a characteristic lifetime analogous to the half-life of a radioactive isotope. Although we will describe these excited states as being energy eigenstates, the fact that they decay tells us that they can't *exactly* be energy eigenstates. It also tells us that there must be some uncertainty as to the exact energy value associated with those states. There will be observational consequences of this, although in practice for real atoms these consequences are extremely difficult to observe.

11.4.2   Vacuum Energy

A second consequence of this formulation of the Heisenberg Uncertainty Principle is the possibility of vacuum energy. Consider a small region of space. Suppose that it's empty; that is, you've taken out everything you can take out of it, including atoms, light (photons), dark matter, and so forth. Make sure that there are no quantum systems anywhere with non-negligible probability for being found in this region of space. Over a finite time interval Δt, you can't be sure exactly how much energy there is in this region of space; your uncertainty in the amount of energy must be at least $\Delta E = \frac{\hbar}{2\Delta t}$. As a result, there *may* be energy in the vacuum.

What is the expectation value of this energy? You might predict that the expectation should be 0, even though the uncertainty has to be greater than zero. Figuring

it out requires going into relativistic quantum mechanics, called quantum field theory. Unfortunately, even quantum field theory can't calculate that right, for naive estimates of what you'd get (the best we can really do) gives a value of the vacuum energy density that is so high that it would prevent galaxies from ever having formed in our Universe. The fact that you are reading this indicates that this estimate cannot be right. Indeed, quantum field theory estimates a value for the vacuum energy density that is 120 orders of magnitude too big! That's pretty far off. As such, we have to say that we don't completely understand the nature of vacuum energy.

What form would this vacuum energy take? We've already seen that in a finite time interval Δt, we can't say with certainty that the vacuum has zero energy. In quantum field theory, it becomes possible to create and destroy particles, as long as you obey all of the conservation laws. For example, two photons can interact and create an electron/positron pair, where a positron is the antimatter partner to an electron. If you don't have to worry about conserving energy, however, you can create a positron/electron pair out of *absolutely nothing*. . . as long as they re-annihilate back to absolutely nothing fast enough. For every fundamental particle that exists, this sort of thing is going on around us all the time.

What is the net energy density of the vacuum as a result of all of this? For a long time, many physicists assumed that a various terms would cancel out to zero; the naive calculations indicated something absurd, and the most natural result if those calculations are wrong is that things would cancel out. However, in the last ten years, observations of the expansion of the Universe have shown that the expansion is accelerating; indeed, these astronomical observations were the source of the 2011 Nobel Prize in Physics. We don't know what is causing this, and have given the name "dark energy" to whatever it is that is causing it. The simplest explanation for dark energy is that it is vacuum energy. Measurements from cosmology indicate a vacuum energy density corresponding to about 10^{-29} grams per cubic centimeter. That is, the energy density of vacuum energy is 29 orders of magnitude less than the mass-energy density of water. Obviously, we can ignore this in our every day life. However, if you look at the Universe as a whole, most of it is empty; our planet is a very special place that is, compared to most of the Universe, extremely dense with regular atoms. In the Universe as a whole, dark energy makes up three quarters of the energy density. Even though this density may be 120 orders of magnitude smaller than what naive estimates from our theory would suggest, it is coming to dominate the evolution of our Universe.