V. The Postulates of Quantum Mechanics

Within physics, no subject is more controversial than the postulates of quantum mechanics. By and large, all (or most) physicists agree on how to perform calculations in quantum mechanics, but they disagree on exactly what the right way to formulate the postulates is. In these lectures, we will give what is termed the Copenhagen interpretations, but this is not to suggest that this is the only or even the preferred interpretation.

A. Revisiting Representations

In the previous chapter, we talked about *observables*, operators whose basis sets are complete, such that any wave function can be written in terms of them. It's time to revisit some of our examples.

Two sets of operators we discussed first were the position and momentum operators **Q** and **P**, defined by

$$\langle \mathbf{r} | \mathbf{Q} | \psi \rangle = \mathbf{r} \langle \mathbf{r} | \psi \rangle \tag{5.1a}$$

$$\langle \mathbf{r} | \mathbf{P} | \psi \rangle = \frac{\hbar}{i} \nabla \langle \mathbf{r} | \psi \rangle \tag{5.1b}$$

If you take the Hermitian conjugate of (5.1b), it is easy to see that $|\mathbf{r}\rangle$ is an eigenstate of all three operators **Q** with eigenvalues given by **r**. If, on the other hand, you substitute the explicit form (4.23) in (5.1b), you will find $|\mathbf{k}\rangle$ is an eigenstate of **P**, so we have

$$\mathbf{Q} |\mathbf{r}\rangle = \mathbf{r} |\mathbf{r}\rangle \tag{5.2a}$$

$$\mathbf{P} |\mathbf{k}\rangle = \hbar \mathbf{k} |\mathbf{k}\rangle \tag{5.2b}$$

Each of $|\mathbf{r}\rangle$ and $|\mathbf{k}\rangle$ form a complete basis, and since **Q** and **P** are Hermitian operators, it follows, by definition, that they are observables. This correspondence of observables and things we can measure will be one of our postulates.

Furthermore, to calculate the probability of getting a particular value of the position or the momentum, we needed to calculate $|\psi(\mathbf{r})|^2 = |\langle \mathbf{r} | \psi \rangle|^2$ or $|\tilde{\psi}(\mathbf{k})|^2 = |\langle \mathbf{k} | \psi \rangle|^2$ respectively. This suggests that we need to develop a basis set for our operators, and calculate the magnitude squared of the inner product of these basis sets with our wave function to get the probability. We also noted that after a measurement of, say, the position, the particle will *be* at the position we measured it, *i.e.*, the wave function must change to become an eigenstate of the operator we are measuring. Finally, when we are *not* performing measurements, we expect the wave function to evolve according to Schrödinger's equation, or something similar. We are now prepared to state the postulates of quantum mechanics.

B. The Postulates of Quantum Mechanics

We now attempt to lay out these postulates, keeping our discussion as general as possible, so that as we consider more complicated situations we will not have to change these postulates, while being specific enough to be useful. We start with:

Postulate 1: The state of a quantum mechanical system at time t can be described as a normalized ket $\left|\Psi(t)\right\rangle$ in a complex vector space with a positive definite inner product.

Note that we have avoided specifying exactly *which* vector space we are in. For example, for a single spinless particle in 3 dimensions, we use the set of square integrable wave functions, but we would need more (or less) complicated vector spaces if we had particles with spin, or multiple particles, or lots of other possibilities. The stuff about a "positive definite inner product" just means that $(\psi,\psi)\geq 0$, with equality only if $\psi=0$.

The next statement we need to make is that the wave function evolves according to the Schrödinger equation.

Postulate 2: When you do not perform a measurement, the state vector evolves according to

$$i\hbar\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle = H(t)\left|\Psi(t)\right\rangle, \tag{5.3}$$

where $H(t)$ is an observable.

We will associate $H(t)$ with the energy of the system, but I don't see such an assumption as a necessary part of the postulates. We will call this operator the *Hamiltonian*. Note we have not assumed any particular form for H; doing so would of necessity require that we decide what vector space we are working in. Furthermore, there will be situations where our Hamiltonian will look radically different than cases we have discussed so far. The necessity of making the Hamiltonian Hermitian will be demonstrated in the next section. Stating that this is how it evolves when you *aren't* looking at it is necessary when you compare this with postulate 5.

Postulate 3: For any quantity that one might measure, there is a corresponding observable A, and the results of that measurement can only be one of the eigenvalues a of A

For example, if you measure the *x*-position Q_x of a particle, you can get any number, because Q_x takes on a continuum of eigenvalues. Later we will encounter examples where only discrete values are possible.

Postulate 4:* *Let $\{|a,n\rangle\}$ be an orthonormal basis of the obserable A,with $A|a,n\rangle = a|a,n\rangle$, and let $|\Psi(t)\rangle$ be the wave function at time t. Then the probability of getting the result a at time t will be given by

$$P(a) = \sum_n \left|\langle a,n|\Psi(t)\rangle\right|^2 \tag{5.4}$$

This is akin to saying the probability density of finding a particle in one dimension at a point *x* is given by $\left|\langle x|\Psi\rangle\right|^2$. The sum on *n* is necessary, because there may be multiple basis states with the same eigenvalue *a*.

We have not yet said that *if* you perform a measurement, *and* you get a specific value, the wave function afterwards *must* have that eigenvalue. This is assured by insisting that the wave function be "reduced" by extracting only those portions that have this eigenvalue.

Postulate 5:* *If the results of a measurement of the observable A at time t yields the result a, the wave function immediately afterwards will be given by

$$|\Psi(t^+)\rangle = \frac{1}{\sqrt{P(a)}} \sum_n |a,n\rangle\langle a,n|\Psi(t)\rangle \tag{5.5}$$

As will be discussed later, this change in the wave function is *very* different from the Schrödinger evolution described by postulate 2. The superscripts + is intended to denote the wave function just after the measurement, which is assumed to take zero time. It is easily demonstrated that if you measure *A* again, you will end up getting the result *a* again. Indeed, your wave function won't change.

The postulates, as presented here, presume that the eigenstates of our operator are discrete, rather than continuous, and that the other parameters *n* that describes our basis are discrete. Either or both may be false. If the other parameter *n* is continuous, this doesn't really present much of a problem. You simply define your basis functions in a suitable way; for example, define them as $|a,\nu\rangle$, and the sums over *n* in (5.4) and (5.5) will be converted to integrals over ν. If it bothers you that the form of the postulates depends on whether one is using a discrete or continuum basis, suffice it to say a common notation can be developed which will allow one to work with the two cases as if they were identical.[1]

If the observable we are measuring has a continuum of eigenvalues, then we need to rethink a couple of the postulates. Suppose, for example, that *A* has a continuum of eigenvalues α, with basis states $|\alpha,\nu\rangle$. These states are assumed to be normalized as

$$\langle\alpha',\nu'|\alpha,\nu\rangle = \delta(\alpha'-\alpha)\delta(\nu'-\nu) \tag{5.6}$$

Then the probability of getting *exactly* a particular value α will automatically be zero, because with an infinite number of closely spaced possibilities, the probability of getting

---

[1] Our textbook by Ballantine discusses these mathematical details more extensively.

one of them vanishes. However, we can still calculate the probability of getting α within a specific range. This modifies postulate 4.

Postulate 4b: Let $\{|\alpha,v\rangle\}$ ***be an orthonormal basis of the obserable*** A***, and let*** $|\Psi(t)\rangle$ ***be the wave function at time*** t ***. Then the probability of getting the result*** a ***in the range*** $a_1 < a < a_2$ ***at time*** t ***will be given by***

$$P(\alpha_1 < \alpha < \alpha_2) = \int_{\alpha_1}^{\alpha_2} d\alpha \int dv \left|\langle \alpha, v | \Psi(t) \rangle\right|^2 \tag{5.7}$$

This works well for the fourth postulate, but what do we do with the last postulate? Suppose we measure the observable *A* and discover that the value is α. The problem is that the probability of getting this value is zero, because the integral (5.7) vanishes if we set the two limits on α equal. The key to resolving this problem is to recognize that *all* measurements of physical quantities represented by a continuous variable have errors.[2] Thus, in fact, any measurement of a continuous variable doesn't tell you α, you only learn that α lies in some range, $\alpha_1 < \alpha < \alpha_2$.

Postulate 5b: If the results of a measurement of the observable A ***at time*** t ***yields the continuous result*** α ***measured to lie in the range*** $\alpha_1 < \alpha < \alpha_2$***, the wave function immediately afterwards will be given by***

$$|\Psi(t^+)\rangle = \frac{1}{\sqrt{P(\alpha_1 < \alpha < \alpha_2)}} \int_{\alpha_1}^{\alpha_2} d\alpha \int dv \, |\alpha, v\rangle \langle \alpha, v | \Psi(t) \rangle \tag{5.8}$$

Though we ended up with two versions of the third and fourth postulates, depending on the situation, we will actually always use the first version when performing a general proof. Repeating the proof with the second version tends to be tedious and provide no great insight.

For a single particle, we already know a prescription for finding the Hamiltonian. Specifically, write out the energy as a function of the position **r** and the momentum **p**, and then promote these two classical quantities to the quantum mechanical operators **Q** and **P** respectively. For example,

$$E = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}, t) \quad \to \quad H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{Q}, t) \tag{5.9}$$

It should be emphasized that there is no *necessity* of describing a prescription for turning a classical theory into a quantum mechanical theory. Classical and quantum theories yield experimentally different descriptions of the universe, and the *only* relevant question is whether the observations will agree with one or the other theory. Hence the real question should not be, how do you find the Hamiltonian, but rather, is there a Hamiltonian that we can find that matches the universe we see.

---

[2] Note that this is not necessarily true of discrete variables. I have two children, exactly two, with no error.

It should be noted that the postulates of quantum mechanics may look radically different depending on who you ask. The main problems seem to do with postulates 3, 4 and 5, which have to do with measurement. In particular, the concept of "measurement" is never exactly defined, though we sort of have a sense of taking information "out there" in the universe and getting the information into our measuring devices (or brains, or published papers, or so on). Note that these postulates involve probability, rather than the cool certainty of Newtonian mechanics. We also seem to have two different rules regarding how the wave function changes, postulates 2 and 5, which are radically different from each other. For these reasons, many have tried to rewrite these postulates in a variety of forms, and with varying success.

C. Consistency of the Postulates

The postulates contain a variety of statements that might or might not be consistent with each other. For example, postulate 1 states that the state vector must be normalized, but postulates 2 and 5 describe how the state vector changes without mentioning normalization. Will these two postulates be consistent with normalization? For example, consider Schrödinger's equation. It is helpful to take the Hermitian conjugate of (5.3), which yields the equation

$$-i\hbar\frac{\partial}{\partial t}\langle\Psi(t)| = \langle\Psi(t)|H^\dagger(t) \tag{5.10}$$

where we have, for the moment, deliberately forgotten that the Hamiltonian is Hermitian. Now (5.3) and (5.10) together allow us to calculate the time derivative of the norm squared of the wave function.

$$\begin{aligned}\frac{\partial}{\partial t}\langle\Psi(t)|\Psi(t)\rangle &= \left[\frac{\partial}{\partial t}\langle\Psi(t)|\right]|\Psi(t)\rangle + \langle\Psi(t)|\frac{\partial}{\partial t}|\Psi(t)\rangle \\ &= \frac{1}{i\hbar}\langle\Psi(t)|H^\dagger(t)|\Psi(t)\rangle - \frac{1}{i\hbar}\langle\Psi(t)|H(t)|\Psi(t)\rangle \\ &= \frac{1}{i\hbar}\langle\Psi(t)|\left[H^\dagger(t)-H(t)\right]|\Psi(t)\rangle = 0\end{aligned} \tag{5.11}$$

Therefore $\langle\Psi(t)|\Psi(t)\rangle$ is constant, and therefore will remain nornmalized.

Similarly, if we perform a measurement, then (5.5) describes how the wave function changes. But is it normalized? We use (5.4) as well to find

$$\begin{aligned}\langle\Psi(t^+)|\Psi(t^+)\rangle &= \frac{1}{P(a)}\left\|\sum_n |a,n\rangle\langle a,n|\Psi(t)\rangle\right\|^2 \\ &= P(a)^{-1}\sum_n\sum_m \langle\Psi(t^-)|a,m\rangle\langle a,m|a,n\rangle\langle a,n|\Psi(t)\rangle \\ &= P(a)^{-1}\sum_n\sum_m \langle\Psi(t^-)|a,m\rangle\delta_{mn}\langle a,n|\Psi(t)\rangle\end{aligned}$$

$$\left\langle \Psi\left(t^+\right) \middle| \Psi\left(t^+\right) \right\rangle = P(a)^{-1} \sum_n \left\langle \Psi\left(t^-\right) \middle| a,n \right\rangle \left\langle a,n \middle| \Psi(t) \right\rangle = P(a)^{-1} P(a) = 1 \qquad (5.12)$$

Hence these rules are consistent.

We also have a statement that the outcome has a certain probability. The probability of a *particular* event must always be a positive number, and the sum of all probabilities of every possible outcome must always equal 1. It is obvious from (5.4) that probability is positive, but do they add to one? This is not hard to prove using the completeness relation:

$$\begin{aligned} \sum_a P(a) &= \sum_a \sum_n \left| \left\langle a,n \middle| \Psi(t) \right\rangle \right|^2 = \sum_{a,n} \left\langle \Psi(t) \middle| a,n \right\rangle \left\langle a,n \middle| \Psi(t) \right\rangle \\ &= \left\langle \Psi(t) \middle| \Psi(t) \right\rangle = 1 \end{aligned} \qquad (5.13)$$

Another concern has to do with our choice of basis. If our observable *A* has only non-degenerate eigenvalues, then our eigenstates $\left| a \right\rangle$ are defined up to a phase, but what if we have multiple states $\left| a,n \right\rangle$ which share the same eigenvalue *a*? The specific choice of eigenstates appear in postulates 4 and 5. For example, let's imagine that we have two choices of eigenvectors $\left| a,n \right\rangle$, and $\left| a,n' \right\rangle$. If we calculate the probability in the primed basis, then postulate 4 will look like

$$P'(a) = \sum_{n'} \left\langle \Psi(t) \middle| a,n' \right\rangle \left\langle a,n' \middle| \Psi(t) \right\rangle \qquad (5.13)$$

We now use the completeness relation (4.55) twice to rewrite this as

$$P'(a) = \sum_{a_1,n_1} \sum_{a_2,n_2} \sum_n \left\langle \Psi(t) \middle| a_1,n_1 \right\rangle \left\langle a_1,n_1 \middle| a,n' \right\rangle \left\langle a,n' \middle| a_2,n_2 \right\rangle \left\langle a_2,n_2 \middle| \Psi(t) \right\rangle \qquad (5.14)$$

where $\left| a_i,n_i \right\rangle$ represents the unprimed basis. We now change the sums as follows. First, we note that since eigenstates with different eigenvalues are orthogonal, the sums on a_1 and a_2 will vanish except when $a_1 = a_2 = a$, so we have

$$P'(a) = \sum_{n_1} \sum_{n_2} \sum_n \left\langle \Psi(t) \middle| a,n_1 \right\rangle \left\langle a,n_1 \middle| a,n' \right\rangle \left\langle a,n' \middle| a,n_2 \right\rangle \left\langle a,n_2 \middle| \Psi(t) \right\rangle \qquad (5.15)$$

We now do the same trick in reverse. For the sum over the states $\left| a,n' \right\rangle$, note that we can add additional values of *a*, which we call *a'*, with impunity, since they will not contribute to the sum. So we have

$$P'(a) = \sum_{n_1} \sum_{n_2} \sum_{a',n'} \left\langle \Psi(t) \middle| a,n_1 \right\rangle \left\langle a,n_1 \middle| a',n' \right\rangle \left\langle a',n' \middle| a,n_2 \right\rangle \left\langle a,n_2 \middle| \Psi(t) \right\rangle \qquad (5.16)$$

We then use completeness of the basis $\left| a',n' \right\rangle$ followed by orthonormality of the basis $\left| a,n_i \right\rangle$.

$$\begin{aligned} P'(a) &= \sum_{n_1}\sum_{n_2}\langle\Psi(t)|a,n_1\rangle\langle a,n_1|a,n_2\rangle\langle a,n_2|\Psi(t)\rangle \\ &= \sum_{n_1}\sum_{n_2}\langle\Psi(t)|a,n_1\rangle\delta_{n_1,n_2}\langle a,n_2|\Psi(t)\rangle = \sum_{n}\langle\Psi(t)|a,n\rangle\langle a,n|\Psi(t)\rangle \\ &= \sum_{n}\left|\langle a,n|\Psi(t)\rangle\right|^2 = P(a) \end{aligned} \tag{5.17}$$

Hence the probability is independent of the choice of basis. Similarly, we can show that (5.5) doesn't depend on basis choice.

$$\begin{aligned} \left|\Psi'(t^+)\right\rangle &= \frac{1}{\sqrt{P'(a)}}\sum_{n'}|a,n'\rangle\langle a,n'|\Psi(t)\rangle \\ &= \frac{1}{\sqrt{P(a)}}\sum_{a_1,n_1}\sum_{a_2,n_2}\sum_{n'}|a_1,n_1\rangle\langle a_1,n_1|a,n'\rangle\langle a,n'|a,n_2\rangle\langle a,n_2|\Psi(t)\rangle \\ &= \frac{1}{\sqrt{P(a)}}\sum_{n_1}\sum_{n_2}\sum_{a',n'}|a,n_1\rangle\langle a,n_1|a',n'\rangle\langle a',n'|a,n_2\rangle\langle a,n_2|\Psi(t)\rangle \\ &= \frac{1}{\sqrt{P(a)}}\sum_{n_1}\sum_{n_2}|a,n_1\rangle\langle a,n_1|a,n_2\rangle\langle a,n_2|\Psi(t)\rangle \\ &= \frac{1}{\sqrt{P(a)}}\sum_{n_1}\sum_{n_2}|a,n_1\rangle\delta_{n_1,n_2}\langle a,n_2|\Psi(t)\rangle \\ &= \frac{1}{\sqrt{P(a)}}\sum_{n}|a,n\rangle\langle a,n|\Psi(t)\rangle = \left|\Psi(t^+)\right\rangle \end{aligned} \tag{5.18}$$

Thus the laws of quantum mechanics as laid out in section B are self-consistent, if rather complex and unpalatable. The primary flaw remaining in this formulation of quantum mechanics is the vague concept of measurement, and the distasteful necessity of using probabilities. These are problems we will return to later.

D. Measurements and Collapse of the Wave Function

One of the surprises about quantum mechanics is the way the wave function apparently changes suddenly when you perform a measurement. To illustrate some of the complexities involved, suppose we have a simple two-dimensional spin system. For the moment, consider a system which has two operators which we can measure, which I will call S_x and S_z.[3] In the two-dimensional space, we will assume these take the matrix form

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.19}$$

These are both Hermitian, and as we will soon demonstrate, observables.

---

[3] This is *not* actually an entirely hypothetical situation. The operators are two of the three spin operators **S** for a spin-½ particle such as an electron.

The basis states for S_z are trivial, because the matrix is already diagonalized. The eigenvalues are $\pm\frac{1}{2}\hbar$ and the corresponding eigenvectors are

$$\left|+,z\right\rangle = \begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad \left|-,z\right\rangle = \begin{pmatrix}0\\1\end{pmatrix} \tag{5.20}$$

The basis states for S_x were found around (4.11). The eigenvalues are again $\pm\frac{1}{2}\hbar$ and the corresponding eigenvectors are

$$\left|+,x\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad \text{and} \quad \left|-,x\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \tag{5.21}$$

Now, suppose we start with an initial state

$$\left|\Psi\left(t=0\right)\right\rangle = \begin{pmatrix}1\\0\end{pmatrix} = \left|+,z\right\rangle \tag{5.22}$$

We will now proceed to perform a series of measurements on the state (5.22), in sufficiently rapid succession that there is no time for the state to change due to Schrödinger's equation.

Suppose we measure, say, S_z. S_z has eigenvalues $\pm\frac{1}{2}\hbar$, so by postulate 3, the only result we can get is one of these two values. According to postulate 4, the probability of getting each of these values is given by

$$\begin{aligned} P\left(+\tfrac{1}{2}\hbar\right) &= \left|\left\langle +,z\middle|\Psi\right\rangle\right|^2 = \left|\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right|^2 = \left|1\right|^2 = 1 \\ P\left(-\tfrac{1}{2}\hbar\right) &= \left|\left\langle -,z\middle|\Psi\right\rangle\right|^2 = \left|\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right|^2 = \left|0\right|^2 = 0 \end{aligned} \tag{5.23}$$

So the result is unambiguously that we will get the result $+\frac{1}{2}\hbar$. After the measurement, the wave function, by postulate 5, will be given by

$$\left|\Psi\right\rangle = \left|+,z\right\rangle\left\langle +,z\middle|\Psi\right\rangle \big/ \sqrt{P_+} = \left|+,z\right\rangle \tag{5.24}$$

In other words, it is completely unchanged. You can repeat the measurement as many times as you want, without changing the outcome. Note we have abbreviated $P\left(+\frac{1}{2}\hbar\right)$ as P_+ for brevity.

Now suppose you measure S_x instead. According to postulate 4, the probability of getting different outcomes will be

$$\begin{aligned} P\left(+\tfrac{1}{2}\hbar\right) &= \left|\left\langle +,x\middle|\Psi\right\rangle\right|^2 = \left|\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right|^2 = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \\ P\left(-\tfrac{1}{2}\hbar\right) &= \left|\left\langle -,x\middle|\Psi\right\rangle\right|^2 = \left|\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right|^2 = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \end{aligned} \tag{5.25}$$

This time we can't predict which outcome will occur. Furthermore, before we move on to postulate 5, we need to know which of the two outcomes actually occurs. Depending on which case we have, after the measurement the state will be

$$\begin{aligned} |\Psi\rangle_+ &= |+,x\rangle\langle +,x|\Psi\rangle \big/ \sqrt{P_+} = |+,x\rangle \frac{1/\sqrt{2}}{1/\sqrt{2}} = |+,x\rangle \\ |\Psi\rangle_- &= |-,x\rangle\langle -,x|\Psi\rangle \big/ \sqrt{P_-} = |-,x\rangle \frac{1/\sqrt{2}}{1/\sqrt{2}} = |-,x\rangle \end{aligned} \tag{5.26}$$

So we see in this case the wave function *has* changed as a result of the measurement. If you now perform further measurements of S_x, it is not hard to show that you will continue to get the same result again; *i.e.*, the result is no longer probabilistic, instead it will always come out to the same S_x value it did before. Let's say, for example, that we got the minus value, so the wave function is now $|\Psi\rangle = |-,x\rangle$.

Now switch back and measure S_z again. The probabilities for each of the possible outcomes will now be

$$\begin{aligned} P\left(+\tfrac{1}{2}\hbar\right) &= \left|\langle +,z|\Psi\rangle\right|^2 = \left|\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}\right|^2 = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \\ P\left(-\tfrac{1}{2}\hbar\right) &= \left|\langle -,z|\Psi\rangle\right|^2 = \left|\frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}\right|^2 = \left|-\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \end{aligned} \tag{5.27}$$

Recall that initially, this *same* measurement of S_z produced the result $+\tfrac{1}{2}\hbar$ with probability 1, but now the probability is split evenly between the possible outcomes $\pm\tfrac{1}{2}\hbar$. Performing measurement S_x *changed* the wave function, and as a result, the probabilities changed for the measurement S_z. It can then be shown that after this measurement is performed, depending on the outcome, the wave function will take the form $|\Psi\rangle = |+,z\rangle$ or $|\Psi\rangle = -|-,z\rangle$.

Note that after every measurement, the resulting wave function was an eigenstate of the measurement just performed, as demanded by postulate 5; indeed, in almost every case, it ended up in a simple basis vector. This does *not* generally happen. It happened here because when applying postulate 5, we only had one state with each eigenvalue, and hence the sum had only a single term. Since the final state is guaranteed to be normalized, it must be only that one basis function, times an arbitrary phase. Without going through the argument in detail, it can be shown that there is no physically difference between a state $|\Psi\rangle$ and the same state multiplied by a phase $e^{i\theta}|\Psi\rangle$. Hence, i when there is only one basis vector for each eigenvalue, there is no *harm* in stating that the state after a measurement yielding a specific eigenvalue will be the corresponding basis vector. However, if there are multiple basis vectors with the same eigenvalue, then we will generally *not* get a single basis vector, and you must use (5.5) to calculate the new quantum state $|\Psi\rangle$.

Suppose we start over, and consider only doing two measurements, S_x and S_z of the initial wave function (5.22). If you do S_z first, you will definitely get +, but if you do

S_x and then S_z, the result of the S_z measurement will be uncertain. In other words, the order you perform measurements makes a difference. This happens because the two operators S_x and S_z do not commute. It is not hard to show that if they *do* commute, then the order *does not* matter, so **the order in which you perform two measurements matters if the corresponding operators do not commute**. Note that in a similar manner, the order of multiplication of two operators makes a difference if they do not commute (by definition). At the risk of confusing you by saying something that sounds like a truism, order matters if order matters.[4]

You might wonder what happens if you measure two observables at the same time. The short answer is that this must be impossible, if the observables do not commute. If two observables *do* actually commute, there is no reason in principle you can't measure them at the same time, and we say that the observables are simultaneous observables.

E. Expectation Values and Uncertainties

In chapter 2 we talked about the expectation value and uncertainty of the position and momentum. We want to generalize these to arbitrary operators. Let A be any operator. We'll define the expectation value $\bar{A}$ of A as the average value that we get for A if we measure it for some wave function $|\psi\rangle$. Two notations are used, and we will use both of them, and we have

$$\bar{A} \equiv \langle A \rangle \equiv \sum_a aP(a) \tag{5.28}$$

Where a represents the various possible results of the measurement, and $P(a)$ the corresponding probabilities. With the help of postulate 4, we can rewrite this as

$$\bar{A} \equiv \langle A \rangle = \sum_a a \sum_n \left| \langle a,n|\psi\rangle \right|^2 = \sum_a \sum_n a \langle \psi | a,n \rangle \langle a,n|\psi\rangle \tag{5.29}$$

Now, $|a,n\rangle$ is an eigenstate of A with $A|a,n\rangle = a|a,n\rangle$, so we can use this to rewrite (5.29), then use completeness to show

$$\begin{aligned} \bar{A} &\equiv \langle A \rangle = \sum_a \sum_n \langle \psi | A | a,n \rangle \langle a,n|\psi\rangle, \\ \bar{A} &\equiv \langle A \rangle = \langle \psi | A | \psi \rangle \end{aligned} \tag{5.30}$$

Note that (5.30) allows us to find the expectation value of A *without* first writing it in terms of A's basis vectors. Hence in chapter 2, when we were finding expectation values of p, it is not necessary to find the Fourier transform of $\tilde{\psi}(k)$, we can work directly with $\psi(x)$, and just use the operator form of P.

---

[4] To make this clearer: Order matters when performing measurements if order matters when multiplying.

Although many measurements of similarly constructed wave functions $|\psi\rangle$ will yield some average value, there will be some spread, or uncertainty, in this value defined by

$$(\Delta A)^2 \equiv \left\langle \left(A-\bar{A}\right)^2 \right\rangle \tag{5.31}$$

A more useful version of (5.30) can be found by squaring out the expression, and realizing that any constant inside an expectation value can be taken outside, and the expectation value of any constant is just that constant. So we have

$$(\Delta A)^2 \equiv \left\langle A^2 - 2A\bar{A} + \bar{A}^2 \right\rangle = \left\langle A^2 \right\rangle - 2\bar{A}\langle A\rangle + \left\langle \bar{A}^2 \right\rangle = \left\langle A^2 \right\rangle - 2\bar{A}^2 + \bar{A}^2,$$

$$(\Delta A)^2 = \left\langle A^2 \right\rangle - \langle A\rangle^2 \tag{5.32}$$

When discussing uncertainties in chapter 2, we stated without proof that there was a limit on the product of the uncertainty of the position and momentum of the particle. We will now prove not only this, but a similar relationship between *any* two observables that do not commute. Let *A* and *B* be any two operators. Consider the ket defined by

$$\left[\alpha\left(A-\bar{A}\right) + i\beta\left(B-\bar{B}\right)\right]|\psi\rangle \tag{5.33}$$

Where $|\psi\rangle$ represents the state of our system at some time and α and β are real numbers to be specified later. All kets have a non-negative norm-squared, so we have

$$\left\| \left[\alpha\left(A-\bar{A}\right) + i\beta\left(B-\bar{B}\right)\right]|\psi\rangle \right\|^2 \ge 0,$$

$$\langle\psi|\left[\alpha\left(A-\bar{A}\right) - i\beta\left(B-\bar{B}\right)\right]\left[\alpha\left(A-\bar{A}\right) + i\beta\left(B-\bar{B}\right)\right]|\psi\rangle \ge 0, \tag{5.34}$$

$$\alpha^2 \langle\psi|\left(A-\bar{A}\right)^2|\psi\rangle + \beta^2 \langle\psi|\left(B-\bar{B}\right)^2|\psi\rangle + i\alpha\beta\langle\psi|\left[\left(A-\bar{A}\right),\left(B-\bar{B}\right)\right]|\psi\rangle \ge 0$$

In the last term, notice the presence of a commutator. Since constants commute with anything, the last term simplifies a lot and we have

$$\alpha^2 (\Delta A)^2 + \beta^2 (\Delta B)^2 + \alpha\beta\langle\psi|i[A,B]|\psi\rangle \ge 0 \tag{5.35}$$

We now make the substitution $\alpha = \Delta B$ and $\beta = \mp\Delta A$ in (5.35) to yield two inequalities:

$$2(\Delta A)^2 (\Delta B)^2 \mp (\Delta A)(\Delta B)\langle\psi|i[A,B]|\psi\rangle \ge 0 \tag{5.36}$$

We divide this equation by $(\Delta A)(\Delta B)$, a positive number[5], and discover

$$(\Delta A)(\Delta B) \ge \pm\tfrac{1}{2}\langle\psi|i[A,B]|\psi\rangle \tag{5.37}$$

---

[5] You might be concerned that we can have $(\Delta A)(\Delta B) = 0$, invalidating our proof. Though this is true, one can show this only occurs when $|\psi\rangle$ is an eigenstate of one of our two operators, and in this case, the right hand side of (5.37) vanishes, so that our uncertainty relation becomes the trivial relationship $0 \ge 0$.

The two signs in (5.37) represent two true statements; that is, (5.37) is true with both signs, but whichever one is positive gives the stronger statement, so

$$\left(\Delta A\right)\left(\Delta B\right)\geq \tfrac{1}{2}\left|\left\langle i\left[A,B\right]\right\rangle\right| \tag{5.38}$$

This is the generalized uncertainty principle. For example, let *A* be any of our momentum operators, and *B* any of our position operators, then we have

$$\left(\Delta Q_i\right)\left(\Delta P_j\right)\geq \tfrac{1}{2}\left|\left\langle i\left[Q_i,P_j\right]\right\rangle\right| = \tfrac{1}{2}\left|\left\langle i\left(i\hbar\right)\delta_{ij}\right\rangle\right| = \tfrac{1}{2}\delta_{ij}\left|-\hbar\right| = \tfrac{1}{2}\delta_{ij}\hbar \tag{5.39}$$

In one dimension, we normally write this as $\left(\Delta x\right)\left(\Delta p\right)\geq \tfrac{1}{2}\hbar$, but we are now using operator notation, where the position *x* is associated with the operator Q_x and momentum *p* is associated with the operator P_x. Note, however, that (5.38) applies to *any* two operators, though if those operators commute, it is a trivial relation.

F. Evolution of Expectation Values

Let *A* be any observable that does not explicitly depend on time, though its expectation value might, because the state vector is changing. How will its expectation value change with time due to evolution under Schrödinger's equation? With the help of Schrödinger's equation (5.3) and its Hermitian conjugate (5.10), we find

$$\begin{aligned}\frac{d}{dt}\left\langle A\right\rangle &= \frac{d}{dt}\left\langle \Psi\right|A\left|\Psi\right\rangle = \left(\frac{d}{dt}\left\langle\Psi\right|\right)A\left|\Psi\right\rangle + \left\langle\Psi\right|A\left(\frac{d}{dt}\left|\Psi\right\rangle\right)\\ &= \frac{i}{\hbar}\left(\left\langle\Psi\right|H\left(t\right)\right)A\left|\Psi\right\rangle - \frac{i}{\hbar}\left\langle\Psi\right|AH\left(t\right)\left|\Psi\right\rangle,\\ \frac{d}{dt}\left\langle A\right\rangle &= \frac{i}{\hbar}\left\langle\left[H\left(t\right),A\right]\right\rangle \end{aligned} \tag{5.40}$$

In particular, for any observable *A* that commutes with the Hamiltonian, (5.40) tells us that the expectation value does not change. In particular, if the Hamiltonian itself is time independent, it will obviously commute with itself, and we have

$$\frac{d}{dt}\left\langle H\right\rangle = 0 \tag{5.41}$$

Let's apply this in a particular case. Suppose that we are dealing with a Hamiltonian of the form (4.51), reproduced here.

$$H = \frac{\mathbf{P}^2}{2m} + V\left(\mathbf{Q}\right) \tag{5.42}$$

We would like to know how the expectation value of the six operators **Q** and **P** change as a function of time. We have

$$\frac{d}{dt}\left\langle\mathbf{Q}\right\rangle = \frac{i}{\hbar}\left\langle\left[\mathbf{P}^2/2m + V\left(\mathbf{Q}\right),\mathbf{Q}\right]\right\rangle = \frac{i}{2m\hbar}\left\langle\left[\mathbf{P}^2,\mathbf{Q}\right]\right\rangle = \frac{i}{2m\hbar}\left\langle -2i\hbar\mathbf{P}\right\rangle = \frac{1}{m}\left\langle\mathbf{P}\right\rangle \tag{5.43}$$

where we have used (4.67) to simplify the commutator. For momentum we have

$$\frac{d}{dt}\langle\mathbf{P}\rangle = \frac{i}{\hbar}\left\langle\left[\mathbf{P}^2/2m + V(\mathbf{Q}),\mathbf{P}\right]\right\rangle = \frac{i}{\hbar}\left\langle\left[V(\mathbf{Q}),\mathbf{P}\right]\right\rangle \tag{5.44}$$

To work out this commutator, it is easiest to imagine it acting on an arbitrary wave function.

$$\begin{aligned}\left[V(\mathbf{Q}),\mathbf{P}\right]\psi(\mathbf{r}) &= V(\mathbf{Q})\mathbf{P}\Psi(\mathbf{r}) - \mathbf{P}V(\mathbf{Q})\psi(\mathbf{r}) \\ &= \frac{\hbar}{i}\left\{V(\mathbf{r})\nabla\left[\psi(\mathbf{r})\right] - \nabla\left[V(\mathbf{r})\psi(\mathbf{r})\right]\right\} \\ &= -\frac{\hbar}{i}\left[\nabla V(\mathbf{r})\right]\psi(\mathbf{r}) = -\frac{\hbar}{i}\left[\nabla V(\mathbf{Q})\right]\psi(\mathbf{r}), \\ \left[V(\mathbf{Q}),\mathbf{P}\right] &= \hbar i\nabla V(\mathbf{Q})\end{aligned} \tag{5.45}$$

Plugging this into (5.44), and putting it together with (5.45), we have Ehrenfest's Theorem:

$$\frac{d}{dt}\langle\mathbf{Q}\rangle = \frac{1}{m}\langle\mathbf{P}\rangle \tag{5.46a}$$

$$\frac{d}{dt}\langle\mathbf{P}\rangle = \langle -\nabla V(\mathbf{Q})\rangle \tag{5.46b}$$

These equations are easy to understand classically. The first says that the rate of change of position is momentum over mass, which makes perfect sense. The second says that the rate of change of momentum is equal to minus the gradient of the potential, which is the force. However, in quantum mechanics, a particle does not have a definite position, or a definite momentum, so we have to take expectation values of the position, momentum, and force.

Let's solve a simple problem using Ehrenfest's theorem. Suppose we have a particle in a one-dimensional harmonic oscillator, given by[6]

$$E = \frac{p^2}{2m} + \tfrac{1}{2}kx^2 \tag{5.47}$$

We change this to the quantum Hamiltonian

$$H = \frac{P^2}{2m} + \tfrac{1}{2}kQ^2 \tag{5.48}$$

How will the expectation value of the position and momentum change over time? Using (5.46a) and (5.46b), these will evolve according to

---

[6] Please note that problems like this would commonly be given simply as $H = p^2/2m + \frac{1}{2}kx^2$. Depending on the notation of the author, it must be understood by the reader that *p* and *x* are to be promoted to operators (which we call *P* and *Q*), but many authors incautiously do not distinguish between operators and their corresponding classical quantities. I will attempt to avoid this confusion, at the risk of making readers confused when they confront such equations from other sources.

$$\frac{d}{dt}\langle Q\rangle = \frac{1}{m}\langle P\rangle \tag{5.49a}$$

$$\frac{d}{dt}\langle P\rangle = -k\langle Q\rangle \tag{5.49b}$$

Plugging the first equation into the second yields

$$\frac{d^2}{dt^2}\langle Q\rangle = -\frac{k}{m}\langle Q\rangle \tag{5.50}$$

This is identical with the classical equation of motion, and has general solution

$$\langle Q\rangle = A\cos(\omega t + \phi) \tag{5.51}$$

where A and ϕ are arbitrary constants and

$$\omega = \sqrt{k/m} \qquad (5.52),$$

the classical angular frequency. Hence, in this case, we were able to solve (5.46) in a straightforward manner. This situation is unusual, however, because we normally do not have $\langle \nabla V(\mathbf{Q})\rangle = \nabla V(\langle \mathbf{Q}\rangle)$, which made this easy.