Wave Mechanics

Relevant sections in text: 1.1 – 1.6, 2.1

The wave function

We will now create a model for the system that we call a "particle in one dimension". To do this we should define states and observables. Let's start with the quantum mechanical notion of "state".

The state of a quantum system determines, and is determined by, the probability distributions for measurements of all possible observables. Section §1.3 in the text has a nice, brief review of probability concepts. You should read it. I will try to introduce the main concepts as we go.

To begin, we need a definition and an interpretation of "probability". A probability is a number between zero and one that expresses the likelihood of the outcome of a particular measurement compared to all possible outcomes of the measurement. Thus, as you all know, the probability that a flipped coin will land "heads" is 1/2. A more precise interpretation of probability which we shall use in this course, known as the "frequency interpretation", is not without pitfalls. We shall adopt it because it is quite close to what is done experimentally.

If P_a is the probability of getting the outcome "a" of some measurement (*e.g.,* heads) out of all possible outcomes of the measurement, then we expect that if the system is repeatedly prepared in the same state and the measurement repeated then the outcome a will be found in a fraction P_a of experiments, provided we do many experiments. The collection of all probabilities as a function of measurement output is the *probability distribution.* For a flipped coin the probability distribution for the observable "heads or tails" is, of course 1/2 for heads and 1/2 for tails. Note that the total probability for getting either heads or tails is just the sum of the two probabilities, namely unity. This reflects a general result: the probability for a collection of possible outcomes to occur is just the sum of the probabilities of each outcome individually. If one considers *all* possible outcomes the probabilities must add up to one. If one considers some subset of possible outcomes the probabilities still add, but they need not add up to one. So, for example, the probability that a roll of a die will come up with a one is 1/6, the probability that it will come up with 1 or 2 is $1/6 + 1/6 = 1/3$ and so on. In a similar vein, you probably know that the probability that, in two tries, a flipped coin will come up heads is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Thus we see another important probability rule: the probability for a sequence of outcomes is the product of each individual probability in the sequence. You can use the rule: the probability for result "a" OR result "b" is $P_a + P_b$. Similarly, the probability for result "a" AND result "b" is $P_a \times P_b$.

In quantum mechanics, there are a variety of physically equivalent mathematical representations of the state of a system. The oldest representation is due to Schrödinger in which one specifies the state of the system at a given time by specifying a *wave function*. This is a *complex-valued* function of position. In our temporarily one-dimensional universe we write

$$\psi = \psi(x) = f(x) + ig(x),$$

where f and g are real functions and x is a real number representing the possible outcomes of a position measurement for the particle at the given time. Sometimes this kind of wave function is called a "coordinate" or "position" wave function.

This representation of the state is designed to make the computation of probability distributions for the position observable particularly simple. In detail, the probability $\rho(x)dx$ for finding the particle to be between x and $x + dx$ is defined to be

$$\rho(x)dx = |\psi(x)|^2 dx.$$

The square of the absolute value of the function ψ,

$$\rho(x) = |\psi(x)|^2,$$

is a positive, real-valued function called the *probability density for position*. Evidently, just knowing the position probability distribution is sufficient to determine the modulus of the wave function.

The probability $P(a, b)$ that the particle is in a region, $x \in [a, b]$, is obtained by adding up the probabilities, that is, integrating the probability density over the region:

$$P(a, b) = \int_a^b |\psi(x)|^2 dx.$$

We assume that the particle always exists somewhere, that is, if we look throughout the universe (or the region of interest) we will eventually detect the particle. In other words, the probability for finding the particle in the universe is unity:*

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1.$$

This relation is described by saying that ψ is *normalized*. You can see that normalization of the wave function captures the idea that the probability obtained by adding up the probabilities for all possible locations of the the particle is unity.

* If the particle is to be confined to a given region – not necessarily the whole real line – then the integral over that region should be unity.

Note that any function whose absolute value-squared can be integrated over all x (or in the region of interest) to get a finite result can be turned into a wavefunction. Such a function is called *normalizable* or *square-integrable*. Let $\phi(x)$ be such a function. We suppose that

$$\int_{-\infty}^{\infty} |\phi(x)|^2 \, dx = C$$

exists, *i.e.,* $C < \infty$. For any square-integrable function ϕ we can define a corresponding wave function ψ by *normalizing* ϕ, this means we define

$$\psi(x) = \frac{1}{\sqrt{C}} \, \phi(x).$$

You can easily check that ψ is normalized.

Example: *Consider the following four functions*

(a) $\phi(x) = \frac{1}{x^4}$

(b) $\phi(x) = \frac{i}{1+|x|}$,

(c) $\phi(x) = e^{-x^2}$.

(d) $\phi(x) = e^{ix}$

Are they normalizable? If so, normalize them.

Solution:

For the function in (a) the integrand $|\phi|^2 = \frac{1}{x^8}$ vanishes rapidly enough as $x \to \pm\infty$ to converge, but the integrand blows up at $x = 0$ rendering the function *un-normalizable.*

For the function in (b), we have

$$\int_{-\infty}^{\infty} |\phi(x)|^2 \, dx = \int_{-\infty}^{\infty} \frac{1}{(1+|x|)^2} \, dx$$

The integrand is bounded for all x and as $x \to \pm\infty$ the integrand is less than a constant times $1/x^2$. Such an integral will converge and the function is square-integrable. Indeed, we have

$$\int_{-\infty}^{\infty} \frac{1}{(1+|x|)^2} \, dx = 2,$$

the normalized wave function is

$$\psi(x) = \frac{1}{\sqrt{2}} \frac{1}{(1+|x|)^2}.$$

The function in (c) is a Gaussian, as is its square. Gaussians are well known to be normalizable. We have

$$\int_{-\infty}^{\infty} |\phi(x)|^2 \, dx = \int_{-\infty}^{\infty} e^{-2x^2} \, dx = \sqrt{\frac{\pi}{2}}.$$

The normalized wave function is then

$$\psi(x) = \left(\frac{2}{\pi}\right)^{1/4} e^{-x^2}.$$

The function in (d) satisfies

$$|\phi(x)| = 1,$$

so there is no way this function is normalizable.

Expectation value of position and functions of position

When we make a measurement of an observable of a physical system – let's denote the observable by L – we normally get some real number λ. Typically, we repeat the experiment many times and take an average in order to take account of fluctuations in the value of the observable introduced by experimental error, extraneous uncontrollable influences, etc. This average is called the *expectation value* of the observable and is denoted by $\langle L \rangle$. If we know the probability distribution $P(\lambda)$ for all possible outcomes of the measurement, say, $\lambda_1, \lambda_2, \ldots$, we can compute the expectation value by adding* up all possible values of the observable with each term in the sum weighted by the fraction of measurements that such values occur (the probability):

$$\langle L \rangle = \sum_i \lambda_i P(\lambda_i). \tag{1}$$

Let's check this. The probability $P(\lambda_i)$ can be interpreted as the number of times n_i that the outcome λ_i is measured over a large number N of repeated experiments, where $\sum_i n_i = N$. The familiar notion of average that we were taught at an early age is

$$\langle L \rangle = \frac{n_1\lambda_1 + n_2\lambda_2 + \ldots}{N}.$$

Since

$$\frac{n_1\lambda_1 + n_2\lambda_2 + \ldots}{N} = \frac{n_1}{N}\lambda_1 + \frac{n_2}{N}\lambda_2 + \ldots = P(\lambda_1)\lambda_1 + P(\lambda_2)\lambda_2 + \ldots = \sum_i P(\lambda_i)\lambda_i,$$

we recover the result (1) above.

* For simplicity I assume the observable only takes a discrete set of values, so we use summations. If the observable were to take a continuous set of values we would use integrals.

Of course, what we get for the expectation value of an observable will depend upon how we run the experiment. The understanding is that (at least ideally) we create the experimental set-up in the same way for each experimental run. We view the experimental "set-up" as determining the state of the system being measured. The expectation value thus depends upon the state of the system.

Let us apply these ideas in the context of a position measurement in the quantum regime. The only new ingredient here is that the position of a particle is usually viewed as a continuous variable x. This just means that the discreteness, if any, of the position of a particle is so small as to be unmeasurable. The continuum limit of a sum is an integral, and we proceed as follows. Suppose the particle is in a state described by a wave function $\psi(x)$. The expectation value of x in this state, denoted by $\langle x \rangle$, is given by

$$\langle x \rangle = \int_{-\infty}^{\infty} x |\psi(x)|^2 \, dx.$$

Thinking of the integral as a continuous form of a summation, we see that we are indeed adding up all possible values of the observable (this is the x in the integrand) weighted by the probability of finding the particle with position between x and $x + dx$ (this is the $|\psi(x)|^2 dx$ in the integrand).

Likewise, given any observable built as a function $f(x)$ of position, the expectation value of $f(x)$ in the state represented by $\psi(x)$ is

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) |\psi(x)|^2 \, dx.$$

The notion of expectation value is extremely important to the physical predictions made by quantum mechanics — it turns out that all physical predictions of quantum mechanics can be phrased in terms of expectation values. To see what I mean, let's construct the observable whose expectation value gives the probability $P(a, b)$ for the particle to be found in the interval (a, b). Consider the function $f(x)$ defined by

$$f(x) = \begin{cases} 1, & \text{if } x \in [a, b]; \\ 0, & \text{otherwise.} \end{cases}$$

View this function as a quantum mechanical observable, as mentioned above. It is easy to see that the expectation value of f is just the probability that x lies in $[a, b]$:

$$\langle f \rangle = \int_{-\infty}^{\infty} f(x) |\psi(x)|^2 \, dx = \int_a^b |\psi(x)|^2 \, dx.$$

A physical interpretation/application of this observable is that it describes the results of a simple laboratory experiment in which we have a detector which can detect particles in

the region $x \in [a, b]$. The detector "clicks" (modeled by $f = 1$) if a particle passes through the indicated region. The detector is silent (modeled by $f = 0$) otherwise. Suppose we know (*e.g.,* via a given experimental set-up) that a particle is in the state given by the wave function $\psi(x)$ at some given time. We repeat this experiment (with the same set-up – the same state ψ) many times. Depending upon the state, the detector may click on some experimental runs and may not click on other runs. We record all the data and numerically average the results by assigning a value 1 to the runs where the detector clicks and 0 when it doesn't. We find that the average over all experimental runs is the fraction of experimental runs in which the detector clicks.

Problem 1.4 in Griffiths' text

At an instant of time, say $t = 0$, the wave function for a particle is given by

$$\Psi(x) = \begin{cases} A\frac{x}{a}, & \text{if } 0 \le x \le a \\ A\frac{b-x}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise.} \end{cases}$$

(a) *Normalize Ψ (by fixing A).*

We must set

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} |\Psi(x)|^2 \, dx = \int_0^a \left|\frac{A}{a}\right|^2 x^2 \, dx + \int_a^b \left|\frac{A}{b-a}\right|^2 |b-x|^2 \, dx \\ &= \frac{1}{3}|A|^2 a + \frac{1}{3}|A|^2 (b-a) \\ &= \frac{1}{3}|A|^2 b \end{aligned}$$

Therefore, for any real number α, setting $A = \sqrt{\frac{3}{b}} e^{i\alpha}$ will normalize the wave function.

(b) *Sketch $\Psi(x)$ as a function of x.* The graph of Ψ is a spike, increasing linearly from zero to its maximum of A at $x = a$, then decreasing linearly to zero at $x = b$.

(c) *Where is the particle most likely to be found at $t = 0$?* Evidently this is a state in which the particle is known to be between a and b with certainty. From (b) it follows that $|\Psi(x)|^2$ will have its maximum at $x = a$, so the particle is most likely to be found there.

(d) *What is the probability for finding the particle to the left of a at $t = 0$.* Let us call this probability P. We have

$$P = \int_{-\infty}^{a} |\Psi(x)|^2 \, dx = \int_0^a \frac{|A|^2}{a^2} x^2 \, dx = \frac{1}{3}|A|^2 a = \frac{a}{b}.$$

In the limiting case $b = a$ we get $P = 1$ as expected. In the limiting case $b = 2a$ we get $1/2$ as expected.

(e) *What is the expectation value of x?* We have

$$\langle x \rangle = \int_{-\infty}^{\infty} x \left|\Psi(x)\right|^2 dx = \int_0^a \left|\frac{A}{a}\right|^2 x^3\, dx + \int_a^b \left|\frac{A}{b-a}\right|^2 x|b-x|^2\, dx = \frac{a}{2} + \frac{b}{4}.$$

In the limiting case of $b = 2a$ we get $\langle x \rangle = a$, as expected. □

States and measurements

Suppose we set up an experiment and measure the position of a particle and get, say, x_0. Where was the particle immediately before the measurement? This is a thorny question. According to the rules of quantum mechanics it has no answer, only a probability distribution for the possible outcomes. The experimental preparation procedure corresponds to assigning a wave function to the particle and this means that, in principle, the position measurement could have resulted in a variety of outcomes with various probabilities. Quantum mechanics does not commit to a pre-existing reality for the position of the particle that is being measured.

Again, return to the scenario where we set up an experiment and measure the position of a particle and get x_0. What if we measure the position again, immediately *after* the first measurement? Empirically, we should get x_0 again, so we incorporate this idea into the rules of the quantum mechanics game. *Immediately* after a measurement of an observable the state of the system must now be updated to yield that outcome with probability one.*

As a simple example, supposed a particle is measured by a detector to be in a region $x \in (a, b)$. Then, irrespective of the state prior to the measurement, the wave function just *after* the measurement is to be of the form

$$\psi(x) = \begin{cases} h(x), & a < x < b \\ 0, & x < a,\ x > b \end{cases}$$

where $h(x)$ is any function such that

$$\int_a^b |h(x)|^2\, dx = 1.$$

In light of the above resutls, and speaking a little more figuratively than is comfortable for me, it appears that the "reality" of a particle's observable features are somehow brought

* I keep saying "immediately" to prevent the appearance of normal sorts of dynamical effects coming from interaction of the particle with its environment.

into being at the time of the measurement! People have contemplated this aspect of quantum theory since its inception. In particular, people (such as Einstein, Podolsky and Rosen) realized it is conceivable that quantum mechanics is correct as far as it goes, but that a more complete theory could actually say with certainty where the particle was immediately before its location was measured. However, according to a brilliant analysis by John Bell and many subsequent experiments it is believed that this is in fact *not* the case! It is possible to test experimentally whether an observable like position has definite "real" value prior to measurement, and experiments say it does not. It is intriguing, to say the least, to contemplate a universe whose "reality" is contingent upon "measurements". I hope to tell you more about this later in the course.

Dynamics: The Schrödinger equation

As we have seen, two of the main ingredients in a physical framework like quantum mechanics are given by states and observables. There is a third ingredient which must now be introduced: the *dynamical law*.

We have introduced a description of the state of a system via the wave function. In physics, we normally view the state of a system as being associated with a given time. For example, we mentioned that the state of a particle in classical mechanics can be specified by the position x and momentum p (or velocity); of course these observables can (and usually do) change in time. In classical mechanics the physical law governing the change in state in time is presented in the form of a differential equation, usually called the *equation of motion* for the state $(x(t), p(t))$ as a function of time. For example, for a particle moving in a force derivable from a potential energy function $V(x,t)$, we have

$$\frac{dx}{dt} = \frac{p}{m}$$
$$\frac{dp}{dt} = -\frac{\partial V}{\partial x}.$$

I chose to write the equations of motion as a pair of first-order differential equations in time* to emphasize the fact that the solution is specified once initial conditions $(x(0), p(0))$–the initial state–is specified. The differential equations of motion specify the *dynamical law* for the state of the system.

In quantum mechanics the role of the dynamical law is played by the *Schrödinger equation*. It determines the state as a function of time via a 1-parameter family of wave functions:

$$\Psi = \Psi(x,t),$$

* These are the Hamilton equations of motion.

where, for each value of t, $\Psi(x,t)$ is a normalized wave function:

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, dx = 1.$$

For a non-relativistic particle with mass m (we shall always use the non-relativistic approximation), the Schrödinger equation is determined by the choice of potential energy function $V(x,t)$. Given the potential energy, the Schrödinger equation is given by

$$i\hbar \frac{\partial \Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x,t)}{\partial x^2} + V(x,t)\Psi(x,t).$$

Here $\hbar$ is a fundamental constant of nature usually called "h-bar" or *Planck's constant*. Numerically,

$$\hbar = 1.05457 \times 10^{-34} J \ s.$$

(Actually, $\hbar$ is Planck's original constant – denoted h – divided by 2π: $\hbar = \frac{h}{2\pi}$.) You can think of the Schrödinger equation as playing the role in quantum mechanics that Newton's second law ($F = ma$) plays in classical mechanics. Let me now list some the most important features of the Schrödinger equation.

The Schrödinger equation (SE) is a *linear* partial differential equation with variable coefficients (thanks to the potential). Linear means that if $\Psi_1(x,t)$ and $\Psi_2(x,t)$ are solutions, then so is

$$\Psi_3(x,t) = a\Psi_1(x,t) + b\Psi_2(x,t),$$

where a and b are any complex constants.

Because the SE is first-order in time, the wave function $\Psi(x,t)$ is uniquely determined by the SE once its initial value $\Psi(x,0)$ is specified. Consequently, as usual, we give the initial state of the system and the dynamical law determines the state for all time. Thus, for example, if the position wave function is given initially, we can solve the SE and find the probability amplitude for position measurements at any other time.

Because the SE also involves space derivatives, we will also have to give spatial boundary conditions to solve the SE. For example, if the particle moves in an infinite domain then the boundary conditions typically have the modulus of the wave function vanishing as $x \to \pm\infty$. Soon we shall consider a variety of examples of solving the initial-boundary value problem for the Schrödinger equation.

Roughly speaking, Planck's constant controls the regime in which quantum corrections to classical mechanics become important. Note that $\hbar$ is very small in typical macroscopic units of Joules and seconds. This corresponds to the fact that quantum effects are typically negligible for macroscopic phenomena. Of course, the fact that Planck's constant is small in SI units does not really mean that the number is unambiguously small. We can make

Planck's constant take on any value we want by a suitable choice of units. There is a valuable lesson here: numerical quantities with units are never "large" or "small" since the numerical value depends upon units. Only dimensionless ratios are unambiguously large or small. So, what dimensionless number is "small" ("large") when classical (quantum) mechanics is correctly describing a system? This is a slightly delicate question. For now we simply mention that the classical approximation is appropriate when the ratio of $\hbar$ to the classical action integral is very small.

There is one more part of the dynamical rule in quantum mechanics that I must remind you about — I already mentioned it earlier. This rule says: If you make a measurement at time t_0 and, say, find the particle at a position x_0, then a measurement of position made immediately afterward must still give x_0 with probability one. This means that after the measurement the wave function $\psi(x,t)$ "collapses" at t_0 to a δ-function $\delta(x-x_0)$. We will have more to say about the δ-function and the interpretation of this "collapse postulate" later. For now, we have two types of time evolution in quantum mechanics. Time evolution in the *absence* of an external interaction which is "measurement" is defined by the Schrödinger equation. An external interaction which determines the value of an observable will "collapse" the wave function into a state in which that observable takes that value with probability one.*

Probability conservation

Recall that the state of the system, represented by the wave function $\psi(x)$ must be normalized:

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, dx = 1.$$

This is necessary for the probability interpretation of the state. Now we have an interesting problem. The SE is supposed to determine a 1-parameter family of states, $\Psi(x,t)$. If the initial state is normalized, will subsequent states also be normalized? This is non-trivial because the normalization condition is *non-linear* and may not be compatible with the linear SE. I will now show you that, with appropriate boundary conditions, the normalization of the wave function is preserved in time. More precisely, I will show that

$$\frac{d}{dt} \int_{-\infty}^{\infty} |\psi(x,t)|^2 \, dx = 0$$

* One might argue that interactions with the environment which constitute "measurement" should be described by the Schrödinger equation just like any other interaction. This is a fair point to make and people have found ways to do just that. One can then aspire to reconcile the "collapse postulate" with the Schrödinger equation itself. The key idea here is usually called *decoherence*. If you search the internet for this term you can read all about it.

when $\Psi(x,t)$ is a solution to the Schrödinger equation. Consequently, if the wave function is normalized at any one time then it will be normalized for all time. One says that "probability is conserved".**

To see this, we derive a continuity equation for the probability density. Consider the time rate of change of the probability density for position at time t:

$$\frac{\partial}{\partial t}(\Psi^*\Psi) = \Psi^*\frac{\partial \Psi}{\partial t} + \Psi\frac{\partial \Psi^*}{\partial t}.$$

If Ψ satisfies the SE we can compute this time rate of change. First, we need (exercise)

$$\frac{\partial \Psi^*}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2 \Psi}{\partial x^2} + \frac{i}{\hbar}V\Psi^*.$$

We now have (exercise)

$$\frac{\partial}{\partial t}(\Psi^*\Psi) = \frac{\partial}{\partial x}\left[\frac{i\hbar}{2m}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \Psi\frac{\partial \Psi^*}{\partial x}\right)\right].$$

We call

$$j = -\frac{i\hbar}{2m}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \Psi\frac{\partial \Psi^*}{\partial x}\right)$$

the *probability current* for the wave function $\Psi(x,t)$. Thus

$$\frac{d}{dt}\int_{-\infty}^{\infty} |\Psi(x,t)|^2\, dx = \left[\frac{i\hbar}{2m}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \Psi\frac{\partial \Psi^*}{\partial x}\right)\right]_{x=-\infty}^{x=\infty},$$

where we have substituted via the probability current and computed the x integral using the fact that the integrand is the x-derivative of $-j$. If $\Psi(x,t)$ is to be normalizable, then $|\Psi(x,t)|$ must vanish as x approaches $\pm\infty$. Let us assume that Ψ is obtained by solving the SE using this boundary condition. Then the endpoint terms at infinity displayed above must vanish and

$$\frac{d}{dt}\int_{-\infty}^{\infty} |\Psi(x,t)|^2\, dx = 0.$$

We conclude that if Ψ is normalized at any one time, it stays normalized for all time provided we solve the SE with asymptotically vanishing boundary conditions.

Let us see why j is called a probability current. We do this via **Problem 1.14** in the text. Consider the probability P_{ab} that $x \in [a,b]$. We have

$$P_{ab} = \int_a^b \rho(x)\, dx = \int_a^b |\Psi|^2\, dx.$$

** The question as to whether probability is conserved in the process of collapse of matter into a black hole and subsequent evaporation is the "black hole information paradox", representing one the principal problems of theoretical physics.

We assume that Ψ solves the SE. Thus the probability density satisfies

$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial x} = 0,$$

where

$$j = -\frac{i\hbar}{2m}\left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x}\right).$$

Then (exercise)

$$\begin{aligned}\frac{dP_{ab}}{dt} &= \int_a^b \frac{\partial \rho}{\partial t}\, dx \\ &= -\int_a^b \frac{\partial j}{\partial x}\, dx \\ &= j(a,t) - j(b,t),\end{aligned}$$

We interpret this as saying that $j(a,t)$ and $j(b,t)$ are the "flow" or "current" of probability through the boundary $x = a$ and $x = b$ of the region $[a,b]$. Thus the change of probability in $[a,b]$ in time is completely controlled by the net flux of probability through the boundary of this region.

The momentum

We now give an introduction to the mathematical representation of the momentum observable for a particle in the context of the wave function description of the state. A complete definition would allow us to compute from any given wave function ψ the probability distribution for momentum. We shall do this later. For now, we consider a simple (but important) feature of this probability distribution. We postulate that the expectation value of momentum, $\langle p \rangle$, in the state ψ is given by

$$\langle p \rangle = \int_{-\infty}^{\infty} \psi^* \frac{\hbar}{i} \frac{\partial \psi}{\partial x}\, dx.$$

It is possible to give detailed justifications for this definition. But I emphasize that one cannot really "prove" this formula; it is part of the rules of quantum mechanics. If this disturbs you, take note that while we define the momentum to be $m\dot{x}$ in classical mechanics, this is not something you can really prove, but rather it is a useful definition. Take the same point of view for our definition above.

Despite what I just wrote, it is worth giving *some* justification for this definition. Given the state $\Psi(x,t)$ as a function of time, derived from the SE, consider the time rate of change

of the expectation value of position (exercise):

$$\begin{aligned}\frac{d\langle x\rangle}{dt} &= \frac{d}{dt}\int_{-\infty}^{\infty} x\rho(x,t)\,dx \\ &= \int_{-\infty}^{\infty} x\frac{\partial\rho}{\partial t}\,dx \\ &= -\int_{-\infty}^{\infty} x\frac{\partial j}{\partial x}\,dx \\ &= -\int_{-\infty}^{\infty} j\,dx + [xj]_{-\infty}^{\infty}\,.\end{aligned}$$

In the last equality I integrated by parts. The endpoint term vanishes since we assume boundary conditions such that j vanishes as $x \to \pm\infty$. Thus we have

$$\frac{d\langle x\rangle}{dt} = -\int_{-\infty}^{\infty} j\,dx.$$

Now consider the integral of j; we can integrate by parts here, too. Integrate by parts in the first term below; you get (good exercise)

$$-\int_{-\infty}^{\infty} j\,dx = \int_{-\infty}^{\infty}\frac{i\hbar}{2m}\left(\psi\frac{\partial\psi^*}{\partial x} - \psi^*\frac{\partial\psi}{\partial x}\right)dx = -\frac{i\hbar}{m}\int_{-\infty}^{\infty}\psi^*\frac{\partial\psi}{\partial x}\,dx.$$

If we define the expectation value of momentum to be mass multiplied by time rate of change of expectation value of position (which is what an experimentalist might well do), then we do indeed have (exercise)

$$\langle p\rangle = \int_{-\infty}^{\infty}\psi^*\frac{\hbar}{i}\frac{\partial\psi}{\partial x}\,dx.$$

Observables as operators

So far I have explained how to compute the expectation value of position (and functions of position) and I have shown how to compute the expectation value of momentum. There is an important organizational principle operating here that is central to the formalism of quantum mechanics. The idea is that the position and momentum observables can be mathematically represented by *linear operators*. Later we will spend considerable time exploring this, for now I will just introduce the basic idea.

Consider a wave function $\psi = \psi(x)$. We can make another function by multiplication by x; for now I will give this new function the slightly cumbersome designation $\hat{x}\psi$, :

$$\psi(x) \longrightarrow (\hat{x}\psi)(x) := x\psi(x). \tag{2}$$

This new function denoted by $\hat{x}\psi$ will in general be neither normalized nor normalizable (exercise). But let us suppose that $\psi(x)$ vanishes fast enough at infinity so that $(\hat{x}\psi)(x)$ is normalizable; for example,

$$\psi(x) = \left(\frac{2}{\pi}\right)^{\frac{1}{4}} e^{-x^2}, \quad (\hat{x}\psi)(x) = \left(\frac{2}{\pi}\right)^{\frac{1}{4}} xe^{-x^2}.$$

(You should definitely check that $\psi(x)$ is normalized and that $\hat{x}\psi$ is indeed normalizable.) We say that the linear operation (2) "represents" the position observable. More explicitly, we say that position is represented on wave functions by the linear operation of "multiplication by x". We denote this operation by $\hat{x}$, and view $\hat{x}\psi$ as the result of $\hat{x}$ operating on ψ. The hat is to remind us that we are defining an operator. You might justifiably feel that this is a little overkill with the notation. When we discuss momentum below the notation will be a little more useful.

More generally, observables obtained as functions $f(x)$ of the position are represented by the operator

$$\psi(x) \longrightarrow (\hat{f}\psi)(x) := f(x)\psi(x).$$

In particular, if $V(x)$ is the potential energy function for a particle, the *potential energy operator* is

$$(\hat{V}\psi)(x) = V(x)\psi(x).$$

Similarly, we say that momentum is represented by a differentiation operation. Given any wave function $\psi(x)$ we define the momentum operator $\hat{p}$ as the linear operation

$$\psi(x) \longrightarrow (\hat{p}\psi)(x) := \frac{\hbar}{i}\frac{d\psi}{dx}.$$

If we consider a 1-parameter family of wave functions $\Psi(x,t)$ representing states at various times we would define the momentum operation using a partial deriviative, holding t fixed:

$$\Psi(x,t) \longrightarrow (\hat{p}\Psi)(x,t) := \frac{\hbar}{i}\frac{\partial\Psi}{\partial x}.$$

We can also use the "momentum = derivative" rule to define observables that are functions of momentum. To be sure, this is a bit more tricky than the analogous situation with the position observable. (How to define $\cos(p)$?) For simple observables things are pretty easy, however. For example, you know that in classical mechanics the kinetic energy T is one-half the square of the momentum divided by the mass. The corresponding *kinetic energy operator* is given by

$$(\hat{T}\psi)(x) := (\frac{1}{2m}\hat{p}^2\psi)(x) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}.$$

More complicated functions of momentum can be defined (via Fourier analysis), but the definition is a little too technical for now.

From this we see that the Schrödinger equation can be viewed as

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi,$$

where

$$\begin{aligned}(\hat{H}\psi)(x) &= (\hat{T}\psi)(x) + (\hat{V}\psi)(x) \\ &= -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x).\end{aligned}$$

The linear operator $\hat{H}$ is known as the *Hamiltonian* operator.* One says that "the Hamiltonian generates time evolution", since the infinitesimal change in the state in time is given by the action of the Hamiltonian operator on the state.

Using the organizational principle "observable $\longleftrightarrow$ operator" we can rewrite our expectation value formulas in a universal way (exercise):

$$\langle x\rangle = \int \psi^*(x)(\hat{x}\psi)(x)\,dx,$$

$$\langle p\rangle = \int \psi^*(x)(\hat{p}\psi)(x)\,dx,$$

$$\langle T\rangle = \int \psi^*(x)(\hat{T}\psi)(x)\,dx,$$

and so on. I have used the hat notation to indicate a linear operator; eventually one gets used to the operator idea and drops the hats when convenient.

I have repeatedly used the word *linear* to characterize the operators defined above. Hopefully the idea of a linear operator is familiar to you from, say, a course in linear algebra. But let me make clear what this word means since it is pretty important. An operator $\hat{A}$ acting on wave functions,

$$\psi(x) \longrightarrow (\hat{A}\psi)(x),$$

is linear if it satisfies

$$(\hat{A}(a\psi + b\phi))(x) = a(\hat{A}\psi)(x) + b(\hat{A}\phi)(x),$$

for any two wave functions ψ and ϕ and for any two constants† a and b. You can easily check that, say, the momentum operator is linear since differentiation distributes across

* You can see that the Hamiltonian operator represents the total energy of the particle. As you may recall from classical mechanics, the energy expressed as a function of coordinates and momentum is often called the Hamiltonian.

† I remind you the constants can be complex.

sums and commutes with multiplication by a constant. Let's test your understanding of the word "linear". Is the following operator linear?

$$(\hat{A}\psi) = 5\psi + 1$$

Finally, I assert that—with one important caveat—we can represent the expectation value of an observable defined as a function $Q(x,p)$ of position and momentum (thus any observable you can think of in classical mechanics) via

$$\langle Q \rangle = \int \psi^*(x) \left[Q(x, \frac{\hbar}{i}\frac{\partial}{\partial x})\psi(x) \right] dx.$$

This formula says to replace position with multiplication by x and momentum by the derivative operator. We will have a chance to get more familiar with this type of formula later. For now, simply note that the Hamiltonian is such an operator, and the expectation value is computed as indicated.

The subtlety with this last definition is that the expression of a function of coordinates and momentum as an operator can be ambiguous. Consider the classical function $Q = xp$. As written, an operator representative is $\hat{x}\hat{p}$, that is

$$(\hat{x}\hat{p})\psi = \frac{\hbar}{i} x \frac{\partial \psi}{\partial x}.$$

But an equally valid classical way of writing the same function is $Q = px$, which we would write as

$$(\hat{p}\hat{x})\psi = \frac{\hbar}{i}\frac{\partial (x\psi)}{\partial x}.$$

These two operators are not the same. To see this, we compare their actions on a wave function $\psi(x)$:

$$\begin{aligned}
(\hat{x}\hat{p}\psi)(x) &= \frac{\hbar}{i} x \frac{\partial \psi}{\partial x}, \\
(\hat{p}\hat{x}\psi)(x) &= \frac{\hbar}{i}\frac{\partial (x\psi)}{\partial x}, \\
&= \frac{\hbar}{i} x \frac{\partial \psi}{\partial x} + \frac{\hbar}{i}\psi \\
&= (\hat{x}\hat{p}\psi)(x) + \frac{\hbar}{i}\psi(x).
\end{aligned}$$

Evidently,

$$(\hat{x}\hat{p}\psi)(x) - (\hat{p}\hat{x}\psi)(x) = i\hbar\psi(x).$$

Since this relation is true for any wave function, we formally denote this relationship as a relationship between operators:

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar\hat{1},$$

where "$\hat{1}$" is the identity operator,* defined by

$$(\hat{1}\psi)(x) = \psi(x).$$

We say that "the operators corresponding to the position and momentum do not *commute*". This mathematical fact means that the quantum mechanical representation of classical observables is not uniquely defined; one has to make a choice (*e.g.,* $\hat{x}\hat{p}$ or $\hat{p}\hat{x}$?), and this has physical consequences. This ambiguity is called the "factor ordering ambiguity". You can think of this issue conceptually as follows. Classical mechanics can be viewed as an approximation of quantum mechanics. Consequently, not all of the rules of quantum mechanics can be deduced from classical mechanics, instead, the converse holds. The representation of observables (as operators) in quantum mechanics is not unambiguously determined from their representation (as real numbers) in classical mechanics. The text is a bit misleading on this point.

The failure of operators representing observables to commute is at the heart of many of the profound physical features of quantum mechanics. Indeed, in the next section I will introduce the uncertainty principle, which arises from precisely this non-commutativity, although I shall not prove that connection until later.

Problem 1.7 in Griffiths' text

Show that

$$\frac{d\langle p\rangle}{dt} = \langle -\frac{\partial V}{\partial x}\rangle,$$

a result known as Ehrenfest's Theorem.

The time dependence of the expectation value of momentum comes from the time dependence of the wave function in

$$\langle p\rangle = \int_{-\infty}^{\infty} \frac{\hbar}{i}\Psi^*(x,t)\frac{\partial}{\partial x}\Psi(x,t)\,dx.$$

We have

$$\frac{d\langle p\rangle}{dt} = \int_{-\infty}^{\infty} \frac{\hbar}{i}\left(\frac{\partial\Psi^*(x,t)}{\partial t}\frac{\partial\Psi(x,t)}{\partial x} + \Psi^*(x,t)\frac{\partial^2\Psi(x,t)}{\partial x\partial t}\right)dx \tag{3}$$

The Schrödinger equation says

$$\frac{\partial\Psi}{\partial t} = \frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi\right),$$

* It is a common (abuse of) notation to suppress the writing of the identity operator, just writing $\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar$.

and hence

$$\frac{\partial^2\Psi}{\partial x\partial t} = \frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^3\Psi}{\partial x^3} + \frac{\partial V}{\partial x}\Psi + V\frac{\partial\Psi}{\partial x}\right), \tag{4}$$

Use

$$\frac{\partial\Psi^*}{\partial t} = -\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\Psi^*}{\partial x^2} + V\Psi^*\right),$$

in the first term of (3) and use (4) in the second term of (3). This may not seem like a good idea, but just try it; some terms will cancel. You will find

$$\frac{d\langle p\rangle}{dt} = \int_{-\infty}^{\infty}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\Psi^*}{\partial x^2}\frac{\partial\Psi(x,t)}{\partial x} + \Psi^*(x,t)\left[\frac{\hbar^2}{2m}\frac{\partial^3\Psi}{\partial x^3} - \frac{\partial V}{\partial x}\Psi\right]\right)dx$$

We can integrate by parts twice in the first term, using the fact that the wave function must vanish as $x \to \pm\infty$ in order to be normalized, to get all x derivatives (three of them) onto $\Psi(x,t)$. This term then cancels the other third derviative term and we get

$$\frac{d\langle p\rangle}{dt} = \int_{-\infty}^{\infty}\Psi^*(x,t)\left(-\frac{\partial V}{\partial x}\right)\Psi(x,t) = \langle -\frac{\partial V}{\partial x}\rangle.$$

□

In Newtonian mechanics we interpret $-\partial V/\partial x$ as the force on the particle. Eherenfest's theorem shows that the Schrödinger equation implies Newton's second law holds in an average sense in quantum mechanics. Insofar as statistical fluctuations from the mean are negligible, we recover this fundamental part of classical mechanics. Such a situation arises with macroscopic systems. *e.g.,* with suitably large masses, energies, *etc.* The quantum realm is one in which one cannot neglect these statistical fluctuations – one has quantum "uncertainty", which is our next topic.

The uncertainty principle

We have seen how to extract some (probabilistic) information about position and momentum observables using the wave function. In general, the state of a quantum mechanical particle is not associated with a specific value of the coordinate or momentum, but rather with a distribution of values with varying probabilities. To be sure, one can imagine a wavefunction that describes a "well-localized" position or momentum. Just imagine a (normalized) function whose absolute value is non-zero only in some small region in position space. Later we will see how to build states which have a well-localized momentum distribution. However, there are limits to how well-localized these observables can both be. Consideration of this issue leads to a first instance of the *uncertainty principle*. Let us briefly look at a simple example of the uncertainty principle. I will defer a systematic treatment until later.

Let us use for the wave function (at a given time)

$$\psi(x) = \frac{1}{\pi^{1/4}\sqrt{d}}\, e^{-x^2/2d^2}.$$

You can check as an exercise that this function is properly normalized. If you graph this function you will find that it is a "bell-shaped curve" — it is a Gaussian centered around $x = 0$. The width of the "bell" is controlled by the *width* d. Physically, this wave function represents a state in which the particle has a relatively small probability for being found at locations farther from the origin than $x = \pm d$. If d is made small enough, this is a state describing a particle with a pretty well-defined position $x = 0$, statistically speaking.

We can characterize the width of the probability distributions for position and momentum — their statistical uncertainty — using the idea of *dispersion*, also known as *variance*. The dispersion in position is defined to be

$$\Delta x^2 = \langle x^2 \rangle - \langle x \rangle^2.$$

It is straightforward to check that the variance is the average of the square of the deviation of x from its average value,

$$\Delta x^2 = \langle (x - \langle x \rangle)^2 \rangle,$$

so that Δx is the standard deviation of the probability distribution for position.

The variance for an observable (*e.g.,* position in this case) is a property which, in general, will change from state to state. As you may recall from a course in statistics, if the variance of an observable vanishes in a particular state then its statistical uncertainty vanishes in that state and one can say that the observable takes the mean value "with certainty" in that state.

For the Gaussian state shown above it is easy to check that the expectation value of x vanishes:

$$\langle x \rangle = \int_{-\infty}^{\infty} \frac{1}{d\sqrt{\pi}}\, e^{-x^2/d^2}\, x\, dx = 0.$$

This follows because there is an equal probability for any values x and $-x$, so the average is zero. The width or "spread" of this Gaussian probability distribution is therefore controlled by the expectation value of the square of position. We get

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} \frac{1}{d\sqrt{\pi}}\, e^{-x^2/d^2}\, x^2\, dx = \frac{d^2}{2}$$

The variance for the Gaussian state is thus simply

$$\Delta x^2 = \frac{d^2}{2}.$$

We see that the dispersion does indeed grow as d grows; it reflects the statistical "uncertainty" in position given the state ψ.

Now let us consider the dispersion in momentum for the state defined by the Gaussian wave function. The expectation value of momentum in the Gaussian state is (exercise)

$$\langle p\rangle = \int_{-\infty}^{\infty} \frac{1}{d\sqrt{\pi}}\, e^{-x^2/d^2}\, (-\frac{x}{d^2})\, dx = 0.$$

Here I used

$$\frac{\hbar}{i}\frac{d}{dx}\left(e^{-x^2/2d^2}\right) = -\frac{x}{d^2}e^{-x^2/2d^2},$$

and then I used the fact that the integral of an odd function over symmetric limits vanishes. Again, although I won't prove it now, this result follows from the fact that the Gaussian (position!) state has equal probability for *momentum* p and $-p$. The dispersion in momentum for this state therefore comes from the expectation value of the square of the momentum. This is not too hard to calculate; you should try it as an exercise. Using

$$\left(\frac{\hbar}{i}\frac{d}{dx}\right)^2\left(e^{-x^2/2d^2}\right) = \hbar^2\left(\frac{1}{d^2}e^{-x^2/2d^2} - \frac{x^2}{d^4}e^{-x^2/2d^2}\right)$$

and our previous types of integrals you will find:

$$\langle p^2\rangle = \frac{\hbar^2}{2d^2}.$$

Hence, for this state,

$$\Delta p^2 = \frac{\hbar^2}{2d^2}.$$

Here we see something interesting: while the dispersion in position increases with increasing d, the momentum dispersion *decreases* with increasing d. Indeed, we have

$$(\Delta x^2)(\Delta p^2) = \frac{\hbar^2}{4}$$

so that in the Gaussian state the product of the momentum and position variances is a constant! This means that if we try to limit the "uncertainty" in the position by adjusting the state with a decreasing value of d, then the "uncertainty" in the momentum must grow, and *vice versa*. Thus, at least for the Gaussian states, it is impossible to determine "with certainty" (probability one) both the momentum and position. What I mean by this is that if you prepare a state with a well-localized position so that many measurements of identically prepared particles give with probability near unity that the position is, say, at $x = 0$, then the momentum will have a very large spread of values (in the same state) with appreciable probability. Of course, the converse holds as well, with position and momentum interchanged. Somehow, the position and momentum observables are *incompatible*.

Everything we just did only applied to a simple family of possible states of the particle (the Gaussian states parametrized by d). Later we shall prove that, no matter the state, for position and momentum the product of the dispersions must satisfy

$$(\Delta x^2)(\Delta p^2) \geq \frac{\hbar^2}{4}.$$

This called an *uncertainty relation*, and the idea that certain observables are incompatible in the above sense is called the "uncertainty principle". We shall also see that uncertainty relations are not confined just to position and momentum observables.

Let us make it clear what the uncertainty principle for position and momentum does *not* say. It is a popular misconception that the uncertainty principle means that "position is uncertain", or "momentum is uncertain", in the sense that one can never measure these observables with arbitrary accuracy. This is false. The "uncertainty" being discussed corresponds to the statistical spread of results you will get by measurements of position and momentum (as accurately as you like) in a series of experiments in which you always have the particle in the same state. One can certainly have states in which the statistical uncertainty in position is arbitrarily small. One can certainly have states in which the statistical uncertainty in momentum is arbitrarily small. One cannot have states in which the uncertainties in position *and* momentum are arbitrarily small.