Lecture 4: Properties of the Wave Function: Probability, Normalization, Measurement, Expectation Values, Wave Packets and the Uncertainty Principle

Interpretation of ψ: Given that we can solve Schrödinger's equation to generate a wave function $\psi(x,t)$, the most natural first question to ask is "What is ψ?" As the wave function for a particle in general is complex, ψ itself cannot be associated with measurements. Instead, in the case of a wave function spread over a range of possible coordinates x, we identity the product of $\psi(x,t)$ with its complex conjugate $\psi*(x,t)$ – this is clearly positive definite – as a *position probability density.*

> $P(x,t)$ = position probability density = $\psi^*(x,t)\psi(x,t) = |\psi(x,t)|^2$
> Thus $dP(x,t) \equiv |\psi(x,t)|^2 dx$ is the probability of finding the particle is a region dx around x if a measurement is made at time t

So it is fine for the amplitude $\psi(x,t)$ to be complex, as it is the norm-squared of the amplitude $|\psi(x,t)|^2$ that is associated with measurements.

And more should be said: This means that if a large number N of identical experiments were set up, each prepared in the same way, and at the same moment t in each experiment a measurement is made to determine whether the particle is in the interval $[a,b]$ – sometimes the answer will be yes, sometimes no – and if the results of those N experiments were averaged, then as $N \to \infty$ that average would converge to

$$P(x \in [a,b]) \equiv \int_a^b |\psi(x,t)|^2 dx$$

Normalization: The wave function $\psi(x,t)$ will in general describe a particle that is somewhere in a region. If we are discussing a free particle, that might be all of space, $-\infty < x < \infty$. Or perhaps the particle is confined within a potential well with infinitely high walls, where the width of the well is a. Then the relevant values of a might be $[-a/2, a/2]$. Regardless of the situation, as the particle must be somewhere in the allowed region, we must require

> Normalization: $\int_D |\psi(x,t)|^2 dx = 1$

where the integral extends over the domain D, e.g., $\int_{-\infty}^{\infty}$ or $\int_{-a/2}^{a/2}$ for the two examples mentioned above. Of course, wave functions $\psi(x,t)$ we obtain from solving differential equations like the

Schrödinger equation have in general some arbitrary normalization. In that case we quantum mechanics have to fix that situation. We compute

$$\int_D |\psi(x,t)|^2 dx = N \quad \text{then form} \quad \psi_N(x,t) \equiv \frac{1}{\sqrt{N}}\psi(x,t)$$

The wave function we use is the normalized one, $\psi_N(x,t)$. We assume below that this step has been carried out.

Now this normalization condition has to hold for every t of interest, so you may have the horrible thought that we will need to compute a distinct normalizations for every point in time. But there is a standard proof that this is not true – that once a wave function is normalized at any one t, it remains normalized under time evolution. The proof utilizes the Schrödinger equation:

$$\frac{d}{dt}\int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = \int_{-\infty}^{\infty} \left[\psi^*(x,t)\frac{\partial\psi(x,t)}{\partial dt} + \frac{\partial\psi^*(x,t)}{\partial dt}\psi(x,t)\right] dt$$

$$\text{But } \frac{\partial\psi(x,t)}{\partial t} = \frac{i\hbar}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} - \frac{i}{\hbar}V(x)\psi(x,t) \quad \text{and} \quad \frac{\partial\psi^*(x,t)}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2\psi^*(x,t)}{\partial x^2} + \frac{i}{\hbar}V(x)\psi^*(x,t)$$

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

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

$$= \frac{i\hbar}{2m}\left[\psi^*(x,t)\frac{\partial\psi(x,t)}{\partial x} - \frac{\partial\psi^*(x,t)}{\partial x}\psi(x,t)\right]\Bigg|_{-\infty}^{\infty} = 0$$

Thus the normalization is constant in time:

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

A wave function normalized at some t then remains normalized at all other t.

Domains and continuous and discrete probability distributions: The examples used above correspond to particles whose coordinate domain D can take any value from an infinite, continuous set, e,.g., $x \in [-\infty, \infty]$ or $x \in [-a/2, a/2]$. Thus the probability normalization involves an integral over a continuous probability distribution, as we wrote above:

$$\int_D |\psi(x,t)|^2 dx = 1$$

Our first focus in QM will be problems like the square well, harmonic oscillator, or the free particle where the question we ask is "what is the probability we will find the particle within dx of x" and as dx is infinitesimal, there are an infinite number of xs and thus we need to integrate over all possibilities to make sure we find the particle. Thus the integral above must be 1, correspondingly. But not all questions in quantum mechanics have an infinite set of answers. Sometimes there are only a few choices. While it will be a while before we talk about spin, an electron has a spin, and if place in a magnetic field, can either point up or down. So we might have no interest in where the electron is, but instead want to know its spin. We could invent something analogous to a wave function in this case, a vector of amplitudes that would include just two possibilities:

$$\vec{v} \equiv \begin{pmatrix} c_\uparrow \\ c_\downarrow \end{pmatrix}$$

Then the normalization condition would become

$$\int_D |\psi(x,t)|^2 dx = 1 \qquad \Rightarrow \qquad |c_\uparrow|^2 + |c_\downarrow|^2 = 1 \quad \Leftrightarrow \quad \sum_{i \in \{\uparrow,\downarrow\}} |c_i|^2 = 1$$

That is, our domain D in the spin cases has only two possibilities.

Griffiths spends quite some time discussing discrete distributions and introducing quantities like means, medians, standard deviations, etc., and as QM is all about amplitudes and probabilities, he's probably right to review this material for those who may be less familiar with the subject. As you may be more familiar with discrete probabilities – heads or tails – let's start by emphasizing that continuous and discrete transition probabilities are really not that different. We will soon consider an electron in an infinite square well of width a, which confines the electron to reside in the x domain $[-a/2, a/2]$. Our probability distribution is continuous

$$\int_{-a/2}^{a/2} |\psi(x,t)|^2 dx = 1$$

But remembering our calculus

$$\int_{-a/2}^{a/2} |\psi(x,t)|^2 dx = \lim_{N\to\infty} \sum_{i=1}^{N} |\psi(x_i,t)|^2 \Delta \quad \text{where} \quad \Delta = \frac{a}{N}$$

We have broken the interval $[-a/2, a/2]$ into N equal bins of width Δ, with the x_i labeling the center-points of the bins. In other words, a continuous probability distribution is just the limiting case of a discrete probability distribution. Comparing to our spin case one sees an analogy

between $|c_i|^2$ and $|\psi(x_i,t)|^2\Delta$, the probability of being in a neighborhood of width Δ centered on x_i.

Whether one has a discrete probability distribution or a continuous one, interesting properties can be determined by evaluating moments of distributions. To give us a concrete example of a discrete probability distribution, consider the midterm scores for a large class with N_T students in which scores can range from 0 to 10. We could add up all the scores, then determine $N[j]$, the number of tests with score j for the 11 possibile *outcomes*. Then we can form a discrete probability distribution

$$P[j] \equiv \frac{N(j)}{N_T} \quad \text{(the fraction of tests with score } j\text{)} \qquad \sum_{j=0}^{10} P(j) = 1$$

So we have created a discrete probability distribution for a problem with a finite set of outcomes j, just as previously we formed the continuous probability distribution

$$|\Psi(x)|^2 dx \qquad \int_D |\psi(x,t)|^2 dx = 1$$

for a problem that had an infinite number of possible outcomes (the locations x, a continuous variable, within our interval D).

Now we can determine properties of a probability distribution (discrete or continuous) by evaluating various moments of the outcome variable. The simplest would be the zeroth moment, defines here as the unweighted moment, which just checks that

$$\langle 1 \rangle \equiv \sum_{j=0}^{10} 1P(j) = 1 \quad \text{and} \quad \langle 1 \rangle \equiv \int_D 1|\psi(x,t)|^2 dx = 1$$

The next is the first moment – called the mean –

$$\langle j \rangle \equiv \sum_{j=0}^{10} jP(j) \quad \text{and} \quad \langle x \rangle \equiv \int_D x|\psi(x,t)|^2 dx$$

If one threw all of the exams in a bin, randomly selected one, and recorded its score, then repeated these three steps many time, the average of the recorded scores is the mean.

The second moment of most interest is the variance, the second moment about the mean

$$\langle j - \langle j \rangle \rangle^2 \equiv \sum_{j=0}^{10} (j - \langle j \rangle)^2 P(j) \quad \text{and} \quad \langle x - \langle x \rangle \rangle^2 \equiv \int_D (x - \langle x \rangle)^2 |\psi(x,t)|^2 dx$$

Now one can quickly see, as $\langle j \rangle$ and $\langle x \rangle$ are just numbers, that

$$\langle j - \langle j \rangle \rangle^2 = \langle j^2 - 2j\langle j \rangle + \langle j \rangle^2 \rangle = \langle j^2 \rangle - \langle j \rangle^2 \geq 0 \quad \text{and} \quad \langle x - \langle x \rangle \rangle^2 = \langle x^2 \rangle - \langle x \rangle^2 \geq 0$$

So knowledge of the first and second moments

$$\langle j^2 \rangle \equiv \sum_{j=0}^{10} j^2 P(j) \quad \text{and} \quad \langle x^2 \rangle \equiv \int_D x^2 |\psi(x,t)|^2 dx$$

allows one to calculate the variance. The standard deviation

$$\sigma \equiv \sqrt{\langle j - \langle j \rangle \rangle^2} \quad \text{or} \quad \sqrt{\langle x - \langle x \rangle \rangle^2}$$

is a measure of the breadth of a distribution. For distributions that are approximately Gaussian, the probability that an outcome will be within 1 σ of the mean is about 2/3rds.

One can continue. For example the skewness of a distribution measures its asymmetry about the mean – whether it tilts left or right. It is defined as

$$\frac{1}{\sigma^{3/2}} \langle j - \langle j \rangle \rangle^3 \quad \text{or} \quad \frac{1}{\sigma^{3/2}} \langle x - \langle x \rangle \rangle^3$$

and by expanding as above can be written in terms of the simple third moment, σ, and the mean.

The point of this is to make you comfortable that continuous distributions are a straight-forward extension of the discrete probability distributions you have may have seen before.

Measurement and collapse of the wave function: A pitcher threw a fastball, and the hitter layed off. The ball was caught by the catcher. But the umpire said nothing. Finally the batter said, "Well what was it, a ball or a strike?" The umpire replied "It ain't nothing til I call it." This is also a pretty good description of the relationship of wave functions to measurement.

The position of an electron might be described by a probability $|\psi(x,t)|^2$ that is nonzero over some range in x, but when a measurement is done, the electron's location will be found to be some definite outcome x_1. If in fact 1000 measurement experiments were prepared, all with the initial conditions set up identically, each of these would give a definite value on making a measurement, yielding $\{x_1, x_2, ..., x_{1000}\}$. The individual experiments would likely all yield different results, but

QM tells us that if we look at the distribution of such events, it will match the distribution computed from $|\phi(x)|^2$. That is – this is called the Copenhagen interpretation of QM – $|\phi(x)|^2$ predicts the probabilities of all the possible outcomes x_i, but does not tell one the specific value x_i that will emerge from a specific measurement.

This implies something interesting about measurement: it impacts the wave function. If one makes a measurement that produces the result x_1 at some time t then repeats the measurement immediately afterwards, at some time $t + \delta t$, then the same outcome $x_2 \sim x_1$ (within some small δx) will be obtained. The first measurement has impacted the wave function, *collapsing* it, greatly narrowing the possibilities.. The electron was described immediately prior to measurement by the broad probability distribution $|\psi(x,t|^2$. But at $t + \delta t$, the electron is now described by a very narrow wave packet $|\psi'(x, t + \delta t)|^2$ centered at $x \sim x_1$. How narrow depends on how precisely the measurement was done, and also by how long the wait δt is. While we will say more below, we have already mentioned the uncertainty principle, the notion that the better one defines x, the broader the spread in p. So this new wave packet, very narrow in x, will contain, as a result of the measurement, many high-momentum components, and these components will spread rapidly as time progresses. The more precise the measurement of x_1, the larger the spread in momentum space components. Thus if you do not make the second measurement immediately, but wait some significant time, you likely will not get an x_2 very near x_1.

This is not classical mechanics, but it is a new theory that is guided by intuitive rules, and they are easy for us to embrace. The ideas are quite beautiful. We have not yet really solved any QM problems up to this point, but that is fine – if you are beginning to intuitively understand how the subatomic world works, that intuition will guide you as you begin to solve problems. QM makes sense!

Expectation values of operators: In our discussion of probability distributions and moments – means and standard deviations – we emphasized their importance in characterizing distribution functions, but the discussion above of measurement now takes us beyond the math, into the physics. If the outcome of our experiment is a particle's position, and if we repeat an experiment with identical initial conditions a thousand times, then we have already how the mean of those position outcomes

relate the the wave function

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

In fact, we can rewrite this statement is a slightly different way that corresponds better to the concept of measurement,

$$\langle \hat{x} \rangle \equiv \int_{-\infty}^{\infty} \psi^*(x,t) \; \hat{x} \; \psi(x,t) \; dx = \int_{-\infty}^{\infty} \psi^*(x,t) \; x \; \psi(x,t) \; dx = \int_{-\infty}^{\infty} x \; |\psi(x,t)|^2 \; dx$$

It may seem like semantics at this point, but think of $\hat{x}$ as an operator that *interrogates* the wave function (the process of taking a measurement), and x as the *outcome* of the interrogation (the experimental result).

The momentum operator provides another example of the distinction drawn above – interrogation vs. outcome. We deduced the momentum operator from our discussion of the Schrödinger equation. So

$$\langle \hat{p} \rangle = \int_{-\infty}^{\infty} \psi^*(x,t) \; \hat{p} \; \psi(x,t) \; dx = \int_{-\infty}^{\infty} \psi^*(x,t) \; \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \; \psi(x,t) \; dx$$

Here we can write out the needed interrogation operator for our coordinate-space wave function, but we can't actually evaluate the outcome if we do not have an explicit form for our wave function. But if someone tells you that $\psi(x,t)$ is a normalized plane wave confined to a 1D "volume" of length L, then you can carry out the operation to obtain an outcome

$$\psi(x,t) = \frac{1}{\sqrt{L}} e^{i(p_0 x - E_0 t)/\hbar} \quad \text{where} \quad E_0 = E_0(p) = p_0^2/2m \Rightarrow$$

$$\int_{-L/2}^{L/2} \psi^*(x,t) \; \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \; \psi(x,t) \; dx = \int_{-L/2}^{L/2} \psi^*(x,t) \; p_0 \; \psi(x,t) \; dx = p_0 \int_{-L/2}^{L/2} |\psi(x,t)|^2 \; dx = p_0$$

I should stess that p_0 is just a number in the work above - a parameter defining the wave function.

Time evolution of operator expectation values: Consider an operator like $\hat{x}$ or $\hat{p}$ that itself does not depend on time, but where the wave function it acts on is evolving in time. This would be the case, for example, of a wave packet moving with some velocity. Operator expectation values would then evolve it time because of the wave function change. Following the same steps we employed to demonstrate the normalizations do not evolve in time, we get

$$\frac{d\langle \hat{x} \rangle}{dt} = \frac{i\hbar}{2m} \int_{-\infty}^{\infty} x \; \frac{\partial}{\partial x} \left[\psi^*(x,t) \frac{\partial \psi(x,t)}{\partial x} - \frac{\partial \psi^*(x,t)}{\partial x} \psi(x,t) \right] dx$$

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

$$= -\frac{i\hbar}{m}\int_{-\infty}^{\infty}\psi^*(x,t)\ \frac{\partial}{\partial x}\ \psi(x,t)dx = \frac{1}{m}\int_{-\infty}^{\infty}\psi^*(x,t)\ \left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\ \psi(x,t)dx$$

where we integrated by parts to get the second line, assuming that the wave packet vanishes at the boundaries; and then integrate by parts again in the third line. Thus we find

$$\frac{d\langle\hat{x}\rangle}{dt} = \frac{\langle\hat{p}\rangle}{m} \equiv \langle\hat{v}\rangle$$

One can repeat the steps above starting with $\frac{d\langle\hat{p}\rangle}{dt}$ to find

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

which we recognize at Newton's second law. These two results constitute

> Ehrenfest's theorm (an example of the correspodence principle):
> expectation values obey the corresponding classical laws of motion

The uncertainty principle; If you have taken linear algebra you may have seen the Cauchy-Schwarz inequality. If $\vec{u}$ and $\vec{v}$ are two vectors in some vector space, then

$$\vec{u}\cdot\vec{u}\ \vec{v}\cdot\vec{v} \geq |\vec{u}\cdot\vec{v}|^2$$

You might be able to see – consider viewing our continuous integrals as discrete sums over small intervals Δ – that the integrals we have been performing over normalized functions, e,.g.,$\int\psi^*(x)\psi(x)dx$, can be viewed as dot products, defined for complex functions that are infinite dimensional. Thus it is not surprising that for functions there is an analogous Cauchy-Schwarz identity

$$\int|f(x)|^2dx\ \int|g(x)|^2dx\ \geq\ \left|\int f^*(x)g(x)dx\right|^2$$

We make the following definitions

$$(\Delta x)^2 \equiv \langle(\hat{x}-\langle\hat{x}\rangle)^2\rangle = \langle\hat{x}^2\rangle - \langle\hat{x}\rangle^2 \qquad (\Delta p)^2 \equiv \langle(\hat{p}-\langle\hat{p}\rangle)^2\rangle = \langle\hat{p}^2\rangle - \langle\hat{p}\rangle^2$$

and the following substitutions into the Cauchy-Schwarz identity

$$f(x) \to (\hat{x}-\langle\hat{x}\rangle)\ \psi(x,t) \quad \text{and} \quad g(x) \to (\hat{p}-\langle\hat{p}\rangle)\ \psi(x,t)$$

The LHS is seen to be

$$(\Delta x)^2 \ (\Delta p)^2$$

while the RHS can be manipulated into the form

$$\frac{1}{4}\left| \int \psi^*(x,t) \left[(\hat{x} - \langle \hat{x} \rangle)(\hat{p} - \langle \hat{p} \rangle) - (\hat{p} - \langle \hat{p} \rangle)(\hat{x} - \langle \hat{x} \rangle) \right] \psi(x,t) dx \right|^2$$

$$+\frac{1}{4}\left| \int \psi^*(x,t) \left[(\hat{x} - \langle \hat{x} \rangle)(\hat{p} - \langle \hat{p} \rangle) + (\hat{p} - \langle \hat{p} \rangle)(\hat{x} - \langle \hat{x} \rangle) \right] \psi(x,t) dx \right|^2$$

$$\geq \frac{1}{4}\left| \int \psi^*(x,t) \left[(\hat{x} - \langle \hat{x} \rangle)(\hat{p} - \langle \hat{p} \rangle) - (\hat{p} - \langle \hat{p} \rangle)(\hat{x} - \langle \hat{x} \rangle) \right] \psi(x,t) dx \right|^2$$

as we have the sum of two positive definite terms. Now by direct evaluation you can show

$$[(\hat{x} - \langle \hat{x} \rangle)(\hat{p} - \langle \hat{p} \rangle) - (\hat{p} - \langle \hat{p} \rangle)(\hat{x} - \langle \hat{x} \rangle)] = i\hbar$$

Thus we obtain

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

So we find

> Uncertainty principle: $\Delta x \ \Delta p \geq \frac{\hbar}{2}$

The equality can only hold if the second (even) term above vanishes. Such *coherent states* have quite interesting properties, but will not be covered in this course.