

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare ocw.mit.edu.

PROFESSOR: Anything lingering and disturbing or bewildering? No? Nothing? All right. OK, so the story so far is basically three postulates. The first is that the configuration of a particle is given by, or described by, a wave function ψ of x . Yeah?

So in particular, just to flesh this out a little more, if we were in 3D, for example-- which we're not. We're currently in our one dimensional tripped out tricycles. In 3D, the wave function would be a function of all three positions x , y and z .

If we had two particles, our wave function would be a function of the position of each particle. x_1 , x_2 , and so on. So we'll go through lots of details and examples later on. But for the most part, we're going to be sticking with single particle in one dimension for the next few weeks.

Now again, I want to emphasize this is our first pass through our definition of quantum mechanics. Once we use the language and the machinery a little bit, we're going to develop a more general, more coherent set of rules or definition of quantum mechanics. But this is our first pass.

Two, the meaning of the wave function is that the norm squared ψ of x , norm squared, it's complex, dx is the probability of finding the particle- There's an n in their. Finding the particle-- in the region between x and x plus dx . So ψ squared itself, norm squared, is the probability density. OK?

And third, the superposition principle. If there are two possible configurations the system can be in, which in quantum mechanics means two different wave functions that could describe the system given ψ_1 and ψ_2 , two wave functions that could describe two different configurations of the system. For example, the particles here or the particles over here. It's also possible to find the system in a superposition of

those two ψ is equal to some arbitrary linear combination $\alpha \psi_1 + \beta \psi_2$ of x . OK?

So some things to note-- so questions about those before we move on? No questions? Nothing? Really? You're going to make me threaten you with something. I know there are questions. This is not trivial stuff. OK.

So some things to note. The first is we want to normalize. We will generally normalize and require that the integral over all possible positions of the probability density ψ of x norm squared is equal to 1. This is just saying that the total probability that we find the particle somewhere had better be one. This is like saying if I know a particle is in one of two boxes, because I've put a particle in one of the boxes. I just don't remember which one. Then the probability that it's in the first box plus probability that it's in the second box must be 100% or one. If it's less, then the particle has simply disappeared. And basic rule, things don't just disappear. So probability should be normalized. And this is our prescription.

So a second thing to note is that all reasonable, or non stupid, functions ψ of x are equally reasonable as wave functions. OK? So this is a very reasonable function. It's nice and smooth. It converges to 0 infinity. It's got all the nice properties you might want. This is also a reasonable function. It's a little annoying, but there it is. And they're both perfectly reasonable as wave functions.

This on the other hand, not so much. So for two reasons. First off, it's discontinuous. And as you're going to show in your problem set, discontinuities are very bad for wave functions. So we need our wave functions to be continuous.

The second is over some domain it's multi valued. There are two different values of the function. That's also bad, because what's the probability? It's the norm squared, but if it two values, two values for the probability, that doesn't make any sense. What's the probability that I'm going to fall over in 10 seconds? Well, it's small, but it's not actually equal to 1% or 3%. It's one of those. Hopefully is much lower than that.

So all reasonable functions are equally reasonable as wave functions. And in particular, what that means is all states corresponding to reasonable wave functions are equally reasonable as physical states. There's no primacy in wave functions or in states.

However, with that said, some wave functions are more equal than others. OK? And this is important, and coming up with a good definition of this is going to be an important challenge for us in the next couple of lectures. So in particular, this wave function has a nice simple interpretation. If I tell you this is ψ of x , then what can you tell me about the particle whose wave function is the ψ of x ? What can you tell me about it? What do you know?

AUDIENCE: [INAUDIBLE].

PROFESSOR: It's here, right? It's not over here. Probability is basically 0. Probability is large. It's pretty much here with this great confidence. What about this guy? Less informative, right? It's less obvious what this wave function is telling me. So some wave functions are more equal in the sense that they have-- i.e. they have simple interpretations.

So for example, this wave function continuing on infinitely, this wave function doesn't tell me where the particle is, but what does it tell me?

AUDIENCE: Momentum.

PROFESSOR: The momentum, exactly. So this is giving me information about the momentum of the particle because it has a well defined wavelength. So this one, I would also say is more equal than this one. They're both perfectly physical, but this one has a simple interpretation. And that's going to be important for us.

Related to that is that any reasonable function ψ of x can be expressed as a superposition of more equal wave functions, or more precisely easily interpretable wave functions. We saw this last time in the Fourier theorem. The Fourier theorem said look, take any wave function-- take any function, but I'm going to interpret in the language of quantum mechanics. Take any wave function which is given by some complex valued function, and it can be expressed as a superposition of plane

waves. $1/2\pi$ in our normalization integral $\int dk \tilde{\psi}(k)$, but this is a set of coefficients. e^{ikx} .

So what are we doing here? We're saying pick a value of k . There's a number associated with it, which is going to be an a magnitude and a phase. And that's the magnitude and phase of a plane wave, e^{ikx} . Now remember that e^{ikx} is equal to $\cos kx + i \sin kx$. Which you should all know, but just to remind you. This is a periodic function. These are periodic functions. So this is a plane wave with a definite wavelength, 2π upon k .

So this is a more equal wave function in the sense that it has a definite wavelength. We know what its momentum is. Its momentum is $\hbar k$. Any function, we're saying, can be expressed as a superposition by summing over all possible values of k , all possible different wavelengths. Any function can be expressed as a superposition of wave functions with a definite momentum. That make sense? Fourier didn't think about it that way, but from quantum mechanics, this is the way we want to think about it. It's just a true statement. It's a mathematical fact. Questions about that?

Similarly, I claim that I can expand the very same function, ψ of x , as a superposition of states, not with definite momentum, but of states with definite position. So what's a state with a definite position?

AUDIENCE: Delta.

PROFESSOR: A delta function, exactly. So I claim that any function ψ of x can be expanded a sum over all states with a definite position. So delta of-- well, what's a state with a definite position? $\delta(x - x_0)$. Delta of x minus x_0 . OK? This goes bing when x_0 is equal to x .

But I want a sum over all possible delta functions. That means all possible positions. That means all possible values of x_0 , dx_0 . And I need some coefficient function here. Well, the coefficient function I'm going to call ψ of x_0 .

So is this true? Is it true that I can take any function and expand it in a superposition of delta functions? Absolutely. Because look at what this equation does. Remember,

delta function is your friend. It's a map from integrals to numbers or functions. So this integral, is an integral over x_0 . Here we have a delta of x minus x_0 . So this basically says the value of this integral is what you get by taking the integrand and replacing x by x_0 . Set x equals x_0 , that's when delta equals 0.

So this is equal to the argument evaluated at x_0 is equal to x . That's your ψ of x . OK? Any arbitrarily ugly function can be expressed either as a superposition of states with definite momentum or a superposition of states with definite position. OK? And this is going to be true. We're going to find this is a general statement that any state can be expressed as a superposition of states with a well defined observable quantity for any observable quantity you want.

So let me give you just a quick little bit of intuition. In 2D, this is a perfectly good vector, right? Now here's a question I want to ask you. Is that a superposition? Yeah. I mean every vector can be written as the sum of other vectors, right? And it can be done in an infinite number of ways, right?

So there's no such thing as a state which is not a superposition. Every vector is a superposition of other vectors. It's a sum of other vector. So in particular we often find it useful to pick a basis and say look, I know what I mean by the vector y , y hat is a unit vector in this direction. I know what I mean by the vector x hat. It's a unit vector in this direction.

And now I can ask, given that these are my natural guys, the guys I want to attend to, is this a superposition of x and y ? Or is it just x or y ? Well, that's a superposition. Whereas x hat itself is not. So this somehow is about finding convenient choice of basis. But any given vector can be expressed as a superposition of some pair of basis vectors or a different pair of basis vectors.

There's nothing hallowed about your choice of basis. There's no God given basis for the universe. We look out in the universe in the Hubble deep field, and you don't see somewhere in the Hubble deep field an arrow going x , right?

So there's no natural choice of basis, but it's sometimes convenient to pick a basis.

This is the direction of the surface of the earth. This is the direction perpendicular to it. So sometimes particular basis sets have particular meanings to us. That's true in vectors. This is along the earth. This is perpendicular to it. This would be slightly strange. Maybe if you're leaning.

And similarly, this is an expansion of a function as a sum, as a superposition of other functions. And you could have done this in any good space of functions. We'll talk about that more. These are particularly natural ones. They're more equal. These are ones with different definite values of position, different definite values of momentum. Everyone cool? Quickly what's the momentum associated to the plane wave e^{ikx} ?

AUDIENCE: [INAUDIBLE].

PROFESSOR: $\hbar k$. Good. So now I want to just quickly run over some concept questions for you. So whip out your clickers. OK, we'll do this verbally. All right, let's try this again. So how would you interpret this wave function?

AUDIENCE: e .

PROFESSOR: Solid. How do you know whether the particle is big or small by looking at the wave function?

AUDIENCE: [INAUDIBLE].

PROFESSOR: All right. Two particles described by a plane wave of the form e^{ikx} . Particle one is a smaller wavelength than particle two. Which particle has a larger momentum? Think about it, but don't say it out loud. And this sort of defeats the purpose of the clicker thing, because now I'm supposed to be able to know without you guys saying anything. So instead of saying it out loud, here's what I'd like you to do. Talk to the person next to you and discuss which one has the larger

AUDIENCE: [CHATTER].

All right. Cool, so which one has the larger momentum?

AUDIENCE: A.

PROFESSOR: How come?

[INTERPOSING VOICES]

PROFESSOR: Right, smaller wavelength. P equals $\hbar k$. k equals 2π over λ . Solid? Smaller wavelength, higher momentum. If it has higher momentum, what do you just intuitively expect to know about its energy? It's probably higher. Are you positive about that? No, you need to know how the energy depends on the momentum, but it's probably higher.

So this is an important little lesson that you probably all know from optics and maybe from core mechanics. Shorter wavelength thing, higher energy. Higher momentum for sure. Usually higher energy as well. Very useful rule of thumb to keep in mind. Indeed, it's particle one. OK next one.

Compared to the wave function ψ of x , it's Fourier transform, $\tilde{\psi}$ of x contains more information, or less, or the same, or something. Don't say it out loud. OK, so how many people know the answer? Awesome. And how many people are not sure. OK, good. So talk to the person next to you and convince them briefly.

All right. So let's vote. A, more information. B, less information. C, same. OK, good you got it. So these are not hard ones. This function, which is a sine wave of length l , 0 outside of that region. Which is closer to true? f has a single well defined wavelength for the most part? It's closer to true. This doesn't have to be exact. f has a single well defined wavelengths. Or f is made up of a wide range of wavelengths? Think it to yourself. Ponder that one for a minute.

OK, now before we get talking about it. Hold on, hold on, hold on. Since we don't have clickers, but I want to pull off the same effect, and we can do this, because it's binary here. I want everyone close your eyes. Just close your eyes, just for a moment. Yeah. Or close the eyes of the person next to you. That's fine. And now and I want you to vote. A is f has a single well defined wavelength. B is f has a wide

range of wavelengths. So how many people think A, a single wavelength? OK. Lower your hands, good. And how many people think B, a wide range of wavelengths? Awesome. So this is exactly what happens when we actually use clickers. It's 50/50. So now you guys need to talk to the person next to you and convince each other of the truth.

AUDIENCE: [CHATTER].

All right, so the volume sort of tones down as people, I think, come to resolution. Close your eyes again. Once more into the breach, my friends. So close your eyes, and now let's vote again. f of x has a single, well defined wavelength. And now f of x is made up of a range of wavelengths? OK. There's a dramatic shift in the field to B, it has a wide range of wavelengths, not a single wavelength. And that is, in fact, the correct answer. OK, so learning happens. That was an empirical test. So does anyone want to defend this view that f is made of a wide range of wavelengths? Sure, bring it.

AUDIENCE: So, the sine wave is an infinite, and it cancels out past minus l over 2 and positive l over 2, which means you need to add a bunch of wavelengths to actually cancel it out there.

PROFESSOR: Awesome, exactly. Exactly. If you only had the thing of a single wavelength, it would continue with a single wavelength all the way out. In fact, there's a nice way to say this. When you have a sine wave, what can you say about it's-- we know that a sine wave is continuous, and it's continuous everywhere, right? It's also differentiable everywhere. Its derivative is continuous and differentiable everywhere, because it's a cosine, right?

So if you take a superposition of sines and cosines, do you ever get a discontinuity? No. Do you ever get something whose derivative is discontinuous? No. So how would you ever reproduce a thing with a discontinuity using sines and cosines? Well, you'd need some infinite sum of sines and cosines where there's some technicality about the infinite limit being singular, because you can't do it a

finite number of sines and cosines.

That function is continuous, but its derivative is discontinuous. Yeah? So it's going to take an infinite number of sines and cosines to reproduce that little kink at the edge. Yeah?

AUDIENCE: So a finite number of sines and cosines doesn't mean finding-- or an infinite number of sines and cosines doesn't mean infinite [? regular ?] sines and cosines, right? Because over a finite region [INAUDIBLE].

PROFESSOR: That's true, but you need arbitrarily-- so let's talk about that. That's an excellent question. That's a very good question. The question here is look, there's two different things you can be talking about. One is arbitrarily large and arbitrarily short wavelengths, so an arbitrary range of wavelengths. And the other is an infinite number. But an infinite number is silly, because there's a continuous variable here k . You got an infinite number of wavelengths between one and 1.2, right? It's continuous. So which one do you mean?

So let's go back to this connection that we got a minute ago from short distance and high momentum. That thing looks like it has one particular wavelength. But I claim, in order to reproduce that as a superposition of states with definite momentum, I need arbitrarily high wavelength.

And why do I need arbitrarily high wavelength modes? Why do we need to arbitrarily high momentum modes? Well, it's because of this. We have a kink. And this feature, what's the length scale of that feature? It's infinitesimally small, which means I'm going to have to-- in order to reproduce that, in order to probe it, I'm going to need a momentum that's arbitrarily large.

So it's really about the range, not just the number. But you need arbitrarily large momentum. To construct or detect an arbitrarily small feature you need arbitrarily large momentum modes. Yeah?

AUDIENCE: Why do you [INAUDIBLE]? Why don't you just say, oh you need an arbitrary small wavelength? Why wouldn't you just phrase that [INAUDIBLE]?

PROFESSOR: I chose to phrase it that way because I want to emphasize and encourage-- I emphasize you to think and encourage you to conflate short distance and large momentum. I want the connection between momentum and the length scale to be something that becomes intuitive to you.

So when I talk about something with short features, I'm going to talk about it as something with large momentum. And that's because in a quantum mechanical system, something with short wavelength is something that carries large momentum. That cool? Great. Good question.

AUDIENCE: So earlier you said that any reasonable wave function, a possible wave function, does that mean they're not supposed to be Fourier transformable?

PROFESSOR: That's usually a condition. Yeah, exactly. We don't quite phrase it that way. And in fact, there's a problem on your problem set that will walk you through what we will mean. What should be true of the Fourier transform in order for this to reasonably function. And among other things-- and your intuition here is exactly right-- among other things, being able to have a Fourier transform where you don't have arbitrarily high momentum modes is going to be an important condition. That's going to turn to be related to the derivative being continuous. That's a very good question. So that's the optional problem 8 on problem set 2. Other questions?

PROFESSOR: Cool, so that's it for the clicker questions. Sorry for the technology fail. So I'm just going to turn this off in disgust. That's really irritating.

So today what I want to start on is pick up on the discussion of the uncertainty principle that we sort of outlined previously. The fact that when we have a wave function with reasonably well defined position corresponding to a particle with reasonably well defined position, it didn't have a reasonably well defined momentum and vice versa. The certainty of the momentum seems to imply lack of knowledge about the position and vice versa.

So in order to do that, we need to define uncertainty. So I need to define for you Δx and Δp . So first I just want to run through what should be totally remedial

probability, but it's always useful to just remember how these basic things work.

So consider a set of people in a room, and I want to plot the number of people with a particular age as a function of the age of possible ages. So let's say we have 16 people, and at 14 we have one, and at 15 we have 1, and at 16 we have 3. And that's 16. And at 20 we have 2. And at 21 we have 4. And at 22 we have 5. And that's it. OK. So 1, 1, 3, 2, 4, 5.

OK, so what's the probability that any given person in this group of 16 has a particular age? I'll call it a . So how do we compute the probability that they have age a ? Well this is easy. It's the number that have age a over the total number.

So note an important thing, an important side note, which is that the sum over all possible ages of the probability that you have age a is equal to 1, because it's just going to be the sum of the number with a particular age over the total number, which is just the sum of the number with any given age.

So here's some questions. So what's the most likely age? If you grabbed one of these people from the room with a giant Erector set, and pull out a person, and let them dangle, and ask them what their age is, what's the most likely they'll have?

AUDIENCE: 22.

PROFESSOR: 22. On the other hand, what's the average age? Well, just by eyeball roughly what do you think it is? So around 19 or 20. It turns out to be 19.2 for this. OK. But if everyone had a little sticker on their lapel that says I'm 14, 15, 16, 20, 21 or 22, how many people have the age 19.2? None, right? So a useful thing is that the average need not be an observable value. This is going to come back to haunt us. Oops, 19.4. That's what I got.

So in particular how did I get the average? I'm going to define some notation. This notation is going to stick with us for the rest of quantum mechanics. The average age, how do I compute it? So we all know this, but let me just be explicit about it. It's the sum over all possible ages of the number of the number of people with that age

times the age divided by the total number of people. OK?

So in this case, I'd go 14, 14, 16, 16, 16, 20, 20, 21, 21, 21, 21, 22, 22, 22, 22, 22.

And so that's all I've written here. But notice that I can write this in a nice way. This is equal to the sum over all possible ages of a times the ratio of N_a to N with a ratio of N_a to n total. That's just the probability that any given person has a probability a . a times probability of a . So the expected value is the sum over all possible values of the value times the probability to get that value. Yeah? This is the same equation, but I'm going to box it. It's a very useful relation.

And so, again, does the average have to be measurable? No, it certainly doesn't. And it usually isn't. So let's ask the same thing for the square of ages. What is the average of a squared? Square the ages. You might say, well, why would I ever care about that? But let's just be explicit about it.

So following the same logic here, the average of a squared, the average value of the square of the ages is, well, I'm going to do exactly the same thing. It's just a squared, right? 14 squared, 15 squared, 16 square, 16 squared, 16 squared. So this is going to give me exactly the same expression. So over a of a squared probability of measuring a .

And more generally, the expected value, or the average value of some function of a is equal-- and this is something you don't usually do-- is equal to the sum over a of f of a , the value of f given a particular value of a , times the probability that you measure that value of a in the first place. It's exactly the same logic as averages. Right, cool.

So here's a quick question. Is a squared equal to the expected value of a squared?

AUDIENCE: No.

PROFESSOR: Right, in general no, not necessarily. So for example, the average value-- suppose we have a Gaussian centered at the origin. So here's a . Now a isn't age, but it's something-- I don't know. You include infants or whatever. It's not age.

Its happiness on a given day. So what's the average value? Meh. Right? Sort of vaguely neutral, right? But on the other hand, if you take a squared, very few people have a squared as zero. Most people have a squared as not a 0 value. And most people are sort of in the middle. Most people are sort of hazy on what the day is.

So in this case, the expected value of a , or the average value of a is 0. The average value of a squared is not equal to 0. Yeah? And that's because the squared has everything positive.

So how do we characterize-- this gives us a useful tool for characterizing the width of a distribution. So here we have a distribution where its average value is 0, but its width is non-zero. And then the expectation value of a squared, the expected value of a squared, is non-zero. So how do we define the width of a distribution? This is going to be like our uncertainty. How happy are you today? Well, I'm not sure. How unsure are you? Well, that should give us a precise measure.

So let me define three things. First the deviation. So the deviation is going to be a minus the average value of a . So this is just take the actual value of a and subtract off the average value of a . So we always get something that's centered at 0. I'm going to write it like this.

Note, by the way, just a convenient thing to note. The average value of a minus its average value. Well, what's the average value of 7?

AUDIENCE: 7.

PROFESSOR: OK, good. So that first term is the average value of a . And that second term is the average value of this number, which is just this number minus a . So this is 0. Yeah? The average value of a number is 0. The average value of this variable is the average value of that variable, but that's 0.

So deviation is not a terribly good thing on average, because on average the deviation is always 0. That's what it means to say this is the average. So the derivation is saying how far is any particular instance from the average. And if you average those deviations, they always give you 0. So this is not a very good

measure of the actual width of the system.

But we can get a nice measure by getting the deviation squared. And let's take the mean of the derivation squared. So the mean of the derivation squared, mean of a minus the average value of a squared. This is what I'm going to call the standard deviation. Which is a little odd, because really you'd want to call it the standard deviation squared. But whatever. We use funny words.

So now what does it mean if the average value of a is 0? It means it's centered at 0, but what does it mean if the standard deviation of a is 0? So if the standard deviation is 0, one then the distribution has no width, right? Because if there was any amplitude away from the average value, then that would give a non-zero strictly positive contribution to this average expectation, and this wouldn't be 0 anymore. So standard deviation is 0, as long as there's no width, which is why the standard deviation is a good useful measure of width or uncertainty.

And just as a note, taking this seriously and taking the square, so standard deviation squared, this is equal to the average value of a squared minus twice a times the average value of a plus average value of a quantity squared. But if you do this out, this is going to be equal to a squared minus 2 average value of a average value of a. That's just minus twice the average value of a quantity squared. And then plus average value of a squared. So this is an alternate way of writing the standard deviation. OK? So we can either write it in this fashion or this fashion. And the notation for this is Δa^2 . OK?

So when I talk about an uncertainty, what I mean is, given my distribution, I compute the standard deviation. And the uncertainty is going to be the square root of the standard deviations squared. OK? So Δa , the words I'm going to use for this is the uncertainty in a given some probability distribution. Different probability distributions are going to give me different Δa 's.

So one thing that's sort of annoying is that when you write Δa , there's nothing in the notation that says which distribution you were talking about. When you have multiple distributions, or multiple possible probability distributions, sometimes it's

useful to just put given the probability distribution p of a . This is not very often used, but sometimes it's very helpful when you're doing calculations just to keep track. Everyone cool with that? Yeah, questions?

AUDIENCE: [INAUDIBLE] Δa^2 , right?

PROFESSOR: Yeah, exactly. Of Δa^2 . Yeah. Other questions? Yeah?

AUDIENCE: So really it should be parentheses [INAUDIBLE].

PROFESSOR: Yeah, it's just this is notation that's used typically, so I didn't put the parentheses around precisely to alert you to the stupidities of this notation. So any other questions? Good.

OK, so let's just do the same thing for continuous variables. Now for continuous variables. I'm just going to write the expressions and just get them out of the way. So the average value of some x , given a probability distribution on x where x is a continuous variable, is going to be equal to the integral. Let's just say x is defined from minus infinity to infinity, which is pretty useful, or pretty typical. $\int dx$ probability distribution of x times x . I shouldn't use curly. I should just use x .

And similarly for x^2 , or more generally, for $f(x)$, the average value of $f(x)$, or the expected value of $f(x)$ given this probability distribution, is going to be equal to the integral $\int dx$ minus infinity to infinity. The probability distribution of x times $f(x)$. In direct analogy to what we had before.

So this is all just mathematics. And we define the uncertainty in x is equal to the expectation value of x^2 minus the expected value of x quantity squared. And this is Δx^2 . If you see me dropping an exponent or a factor of 2, please, please, please tell me. So thank you for that.

All of that is just straight up classical probability theory. And I just want to write this in the notation of quantum mechanics. Given that the system is in a state described by the wave function ψ of x , the average value, the expected value of x , the typical value if you just observe the particle at some moment, is equal to the integral over

all possible values of x . The probability distribution, ψ of x norm squared x .

And similarly, for any function of x , the expected value is going to be equal to the integral dx . The probability distribution, which is given by the norm squared of the wave function times f of x minus infinity to infinity. And same definition for uncertainty.

And again, this notation is really dangerous, because the expected value of x depends on the probability distribution. In a physical system, the expected value of x depends on what the state of the system is, what the wave function is, and this notation doesn't indicate that.

So there are a couple of ways to improve this notation. One of which is-- so this is, again, a sort of side note. One way to improve this notation x is to write the expected value of x in the state ψ , so you write ψ as a subscript. Another notation that will come back-- you'll see why this is a useful notation later in the semester-- is this notation, $\langle x \rangle_\psi$. And we will give meaning to this notation later, but I just want to alert you that it's used throughout books, and it means the same thing as what we're talking about the expected value of x given a particular state ψ . OK? Yeah?

AUDIENCE: To calculate the expected value of momentum do you need to transform the--

PROFESSOR: Excellent question. Excellent, excellent question. OK, so the question is, how do we do the same thing for momentum? If you want to compute the expected value of momentum, what do you have to do? Do you have to do some Fourier transform to the wave function? So this is a question that you're going to answer on the problem set and that we made a guess for last time.

But quickly, let's just think about what it's going to be purely formally. Formally, if we want to know the likely value of the momentum, the likely value the momentum, it's a continuous variable. Just like any other observable variable, we can write as the integral over all possible values of momentum from, let's say, it could be minus infinity to infinity. The probability of having that momentum times momentum, right? Everyone cool with that? This is a tautology, right? This is what you mean by

probability.

But we need to know if we have a quantum mechanical system described by state ψ of x , how do we can get the probability that you measure p ? Do I want to do this now? Yeah, OK I do. And we need a guess. Question mark. We made a guess at the end of last lecture that, in quantum mechanics, this should be $\int_{-\infty}^{\infty} |\tilde{\psi}(p)|^2 dp$ to infinity of the Fourier transform. $\tilde{\psi}(p)$ up to an \hbar factor. $\tilde{\psi}(p)$, the Fourier transform p norm squared.

OK, so we're guessing that the Fourier transform norm squared is equal to the probability of measuring the associated momentum. So that's a guess. That's a guess. And so on your problem set you're going to prove it. OK? So exactly the same logic goes through. It's a very good question, thanks. Other questions? Yeah?

AUDIENCE: Is that p the momentum itself? Or is that the probability?

PROFESSOR: So this is the probability of measuring momentum p . And that's the value p . We're summing over all p 's. This is the probability, and that's actually p . So the Fourier transform is a function of the momentum in the same way that the wave function is a function of the position, right?

So this is a function of the momentum. It's norm squared defines the probability. And then the p on the right is this p , because we're computing the expected value of p , or the average value of p . That make sense? Cool. Yeah?

AUDIENCE: Are we then multiplying by p squared if we're doing all p 's? Because we have the dp times p for each [INAUDIBLE].

PROFESSOR: No. So that's a very good question. So let's go back. Very good question. Let me phrase it in terms of position, because the same question comes up. Thank you for asking that. Look at this. This is weird. I'm going to phrase this as a dimensional analysis question. Tell me if this is the same question as you're asking.

This is a thing with dimensions of what? Length, right? But over on the right hand side, we have a length and a probability, which is a number, and then another

length. That looks like x squared, right? So why are we getting something with dimensions of length, not something with dimensions of length squared? And the answer is this is not a probability. It is a probability density. So it's got units of probability per unit length. So this has dimensions of one over length.

So this quantity, $\int p(x) dx$, tells me the probability, which is a pure number, no dimensions. The probability to find the particle between x and $x + dx$. Cool? So that was our second postulate. $\int \psi^*(x) \psi(x) dx$ is the probability of finding it in this domain. And so what we're doing is we're summing over all such domains the probability times the value. Cool? So this is the difference between discrete, where we didn't have these probability densities, we just had numbers, pure numbers and pure probabilities. Now we have probability densities per unit whatever. Yeah?

AUDIENCE: How do you pronounce the last notation that you wrote?

PROFESSOR: How do you pronounce? Good, that's a good question. The question is, how do we pronounce these things. So this is called the expected value of x , or the average value of x , or most typically in quantum mechanics, the expectation value of x . So you can call it anything you want. This is the same thing. The ψ is just to denote that this is in the state ψ . And it can be pronounced in two ways. You can either say the expectation value of x , or the expectation of x in the state ψ . And this would be pronounced one of two ways. The expectation value of x in the state ψ , or ψ x ψ . Yeah. That's a very good question. But they mean the same thing.

Now, I should emphasize that you can have two ways of describing something that mean the same thing, but they carry different connotations, right? Like have a friend who's a really nice guy. He's a mensch. He's a good guy. And so I could see he's a nice guy, I could say he's [*carinoso* ?], and they mean different things in different languages. It's the same idea, but they have different flavors, right? So whatever your native language is, you've got some analog of this.

This means something in a particular mathematical language for talking about quantum mechanics. And this has a different flavor. It carries different implications, and we'll see what that is later. We haven't got there yet. Yeah?

AUDIENCE: Why is there a double notation of ψ ?

PROFESSOR: Why is there a double notation of ψ ? Yeah, we'll see later. Roughly speaking, it's because in computing this expectation value, there's a ψ squared. And so this is to remind you of that. Other questions? Terminology is one of the most annoying features of quantum mechanics. Yeah?

AUDIENCE: So it seems like this [INAUDIBLE] variance is a really convenient way of doing it. How is it the Heisenberg uncertainty works exactly as it does for this definition of variance.

PROFESSOR: That's a very good question. In order to answer that question, we need to actually work out the Heisenberg uncertainty relation. So the question is, look, this is some choice of uncertainty. You could have chosen some other definition of uncertainty. We could have considered the expectation value of x to the fourth minus x to the fourth and taken the fourth root of that. So why this one?

And one answer is, indeed, the uncertainty relation works out quite nicely. But then I think important to say here is that there are many ways you could construct quantities. This is a convenient one, and we will discover that it has nice properties that we like. There is no God given reason why this had to be the right thing. I can say more, but I don't want to take the time to do it, so ask in office hours. OK, good.

The second part of your question was why does the Heisenberg relation work out nicely in terms of these guys, and we will study that in extraordinary detail. We'll see that. So we're going to derive it twice soon and then later. The later version is better.

So let me work out some examples. Or actually, I'm going to skip the examples in the interest of time. They're in the notes, and so they'll be posted on the web page. By the way, the first 18 lectures of notes are posted. I had a busy night last night.

So let's come back to computing expectation values for momentum. So I want to go back to this and ask a silly-- I want to make some progress towards deriving this

relation. So I want to start over on the definition of the expected value of momentum. And I'd like to do it directly in terms of the wave function.

So how would we do this? So one way of saying this is what's the average value of p . Well, I can phrase this in terms of the wave function the following way. I'm going to sum over all positions dx . Expectation value of x squared from minus infinity to infinity. And then the momentum associated to the value x .

So it's tempting to write something like this down to think maybe there's some p of x . This is a tempting thing to write down. Can we? Are we ever in a position to say intelligently that a particle-- that an electron is both hard and white?

AUDIENCE: No.

PROFESSOR: No, because being hard is a superposition of being black and white, right? Are we ever in a position to say that our particle has a definite position x and correspondingly a definite momentum p . It's not that we don't get too. It's that it doesn't make sense to do so. In general, being in a definite position means being in a superposition of having different values for momentum.

And if you want a sharp way of saying this, look at these relations. They claim that any function can be expressed as a superposition of states with definite momentum, right? Well, among other things a state with definite position, x_0 , can be written as a superposition, $\frac{1}{\sqrt{2\pi}} \int dk \delta(k - k_0) e^{ikx}$. I'll call this $\delta(k - k_0)$.

If you haven't played with delta functions before and you haven't seen this, then you will on the problem set, because we have a problem that works through a great many details. But in particular, it's clear that this is not-- this quantity can't be a delta function of k , because, if it were, this would be just e^{ikx} . And that's definitely not a delta function.

Meanwhile, what can you say about the continuity structure of a delta function. Is it continuous? No. Its derivative isn't continuous. Its second derivative. None of its derivatives are in any way continuous. They're all absolutely horrible, OK? So how many momentum modes am I going to need to superimpose in order to reproduce

a function that has this sort of structure? An infinite number. And it turns out it's going to be an infinite number with the same amplitude, slightly different phase, OK?

So you can never say that you're in a state with definite position and definite momentum. Being in a state with definite position means being in a superposition of being in a superposition. In fact, I'm just going right down the answer here. e^{ikx} . Being in a state with definite position means being in a superposition of states with arbitrary momentum and vice versa. You cannot be in a state with definite position, definite momentum. So this doesn't work.

So what we want is we want some good definition. So this does not work. We want some good definition of p given that we're working with a wave function which is a function of x . What is that good definition of the momentum? We have a couple of hints. So hint the first.

So this is what we're after. Hint the first is that a wave-- we know that given a wave with wave number k , which is equal 2π over λ , is associated, according to de Broglie and according to Davisson-Germer experiments, to a particle-- so having a particle-- a wave, with wave number k or wavelength λ associated particle with momentum p is equal to $\hbar k$. Yeah? But in particular, what is a plane with wavelength λ or wave number k look like? That's e^{ikx} . And if I have a wave, a plane wave e^{ikx} , how do I get $\hbar k$ out of it?

Note the following, the derivative with respect to x . Actually let me do this down here. Note that the derivative with respect to x of e^{ikx} is equal to $ik e^{ikx}$. There's nothing up my sleeves.

So in particular, if I want to get $\hbar k$, I can multiply by \hbar and divide by i . Multiply by \hbar , divide by i , derivative with respect to x e^{ikx} . And this is equal to $\hbar k e^{ikx}$. That's suggestive. And I can write this as $p e^{ikx}$.

So let's quickly check the units. So first off, what are the units of \hbar ? Here's the super easy to remember the units of-- or dimensions of \hbar are. $\Delta x \Delta p$ is

\hbar . OK? If you're ever in doubt, if you just remember, \hbar has units of momentum times length. It's just the easiest way to remember it. You'll never forget it that way.

So if \hbar has units of momentum times length, what are the units of k ? 1 over length. So does this dimensionally make sense? Yeah. Momentum times length divided by length number momentum. Good. So dimensionally we haven't lied yet.

So this makes it tempting to say something like, well, \hbar upon i derivative with respect to x is equal in some-- question mark, quotation mark-- p . Right? So at this point it's just tempting to say, look, trust me, p is \hbar upon i derivative with respect to x . But I don't know about you, but I find that deeply, deeply unsatisfying.

So let me ask the question slightly differently. We've followed the de Broglie relations, and we've been led to the idea that using wave functions that there's some relationship between the momentum, the observable quantity that you measure with sticks, and meters, and stuff, and this operator, this differential operator, \hbar upon i derivative with respect to x . By the way, my notation for $\frac{\partial}{\partial x}$ is the partial derivative with respect to x . Just notation.

So if this is supposed to be true in some sense, what is momentum have to do with a derivative? Momentum is about velocities, which is like derivatives with respect to time, right? Times mass. Mass times derivative with respect to time, velocity. So what does it have to do with the derivative with respect to position?

And this ties into the most beautiful theorem in classical mechanics, which is the Noether's theorem, named after the mathematician who discovered it, Emmy Noether. And just out of curiosity, how many people have seen Noether's theorem in class. Oh that's so sad. That's a sin.

OK, so here's a statement of Noether's theorem, and it underlies an enormous amount of classical mechanics, but also of quantum mechanics. Noether, incidentally, was a mathematician. There's a whole wonderful story about Emmy Noether. Ville went to her and was like, look, I'm trying to understand the notion of

energy. And this guy down the hall, Einstein, he has a theory called general relativity about curved space times and how that has something to do with gravity. But it doesn't make a lot of sense to me, because I don't even know how to define the energy. So how do you define momentum and energy in this guy's crazy theory?

And so Noether, who was a mathematician, did all sorts of beautiful stuff in algebra, looked at the problem and was like I don't even know what it means in classical mechanics. So what is a mean in classical mechanics? So she went back to classical mechanics and, from first principles, came up with a good definition of momentum, which turns out to underlie the modern idea of conserved quantities and symmetries. And it's had enormous far reaching impact, and say her name would praise.

So Noether tells us the following statement, to every symmetry-- and I should say continuous symmetry-- to every symmetry is associated a conserved quantity. OK?

So in particular, what do I mean by symmetry? Well, for example, translations. x goes to x plus some length l . This could be done for arbitrary length l . So for example, translation by this much or translation by that much. These are translations. To every symmetry is associated a conserved quantity.

What symmetry is associated to translations? Conservation of momentum, p dot. Time translations, t goes to t plus capital T . What's a conserved quantity associated with time translational symmetry? Energy, which is time independent.

And rotations. Rotational symmetries. x , as a vector, goes to some rotation times x . What's conserved by virtue of rotational symmetry?

AUDIENCE: Angular momentum.

PROFESSOR: Angular momentum. Rock on. OK So quickly, I'm not going to prove to you Noether's theorem. It's one of the most beautiful and important theorems in physics, and you should all study it.

But let me just convince you quickly that it's true in classical mechanics. And this

was observed long before Noether pointed out why it was true in general. What does it mean to have translational symmetry? It means that, if I do an experiment here and I do it there, I get exactly the same results. I translate the system and nothing changes. Cool? That's what I mean by saying I have a symmetry. You do this thing, and nothing changes.

OK, so imagine I have a particle, a classical particle, and it's moving in some potential. This is u of x , right? And we know what the equations of motion are in classical mechanics from f equals ma p dot is equal to the force, which is minus the gradient of u . Minus the gradient of u . Right? That's f equals ma in terms of the potential.

Now is the gradient of u 0? No. In this case, there's a force. So if I do an experiment here, do I get the same thing as doing my experiment there?

AUDIENCE: No.

PROFESSOR: Certainly not. The [? system ?] is not translationally invariant. The potential breaks that translational symmetry. What potential has translational symmetry?

AUDIENCE: [INAUDIBLE].

PROFESSOR: Yeah, constant. The only potential that has full translational symmetry in one dimension is translation invariant, i.e. constant. OK? What's the force?

AUDIENCE: 0.

PROFESSOR: 0. 0 gradient. So what's p dot? Yep. Noether's theorem. Solid. OK. Less trivial is conservation of energy. I claim and she claims-- and she's right-- that if the system has the same dynamics at one moment and a few moments later and, indeed, any amount of time later, if the laws of physics don't change in time, then there must be a conserved quantity called energy. There must be a conserved quantity. And that's Noether's theorem.

So this is the first step, but this still doesn't tell us what momentum exactly has to do with a derivative with respect to space. We see that there's a relationship between

translations and momentum conservation, but what's the relationship?

So let's do this. I'm going to define an operation called translate by L . And what translate by L does is it takes f of x and it maps it to f of x minus L . So this is a thing that affects the translation. And why do I say that's a translation by L rather than minus L . Well, the point-- if you have some function like this, and it has a peak at 0, then after the translation, the peak is when x is equal to L . OK? So just to get the signs straight.

So define this operation, which takes a function of x and translates it by L , but leaves it otherwise identical. So let's consider how translations behave on functions. And this is really cute. f of x minus L can be written as a Taylor expansion around the point x -- around the point L equals 0.

So let's do Taylor expansion for small L . So this is equal to f of x minus L derivative with respect to x of f of x plus L squared over 2 derivative squared, two derivatives of x , f of x plus dot, dot, dot. Right? I'm just Taylor expanding. Nothing sneaky.

Let's add the next term, actually. Let me do this on a whole new board. All right, so we have translate by L on f of x is equal to f of x minus L is equal to f of x . Now Taylor expanding minus L derivative with respect to x of f plus L squared over 2-- I'm not giving myself enough space. I'm sorry. f of x minus L is equal to f of x minus L with respect to x of f of x plus L squared over 2 to derivatives of x f of x minus L cubed over 6-- we're just Taylor expanding-- cubed with respect to x of f of x and so on. Yeah?

But I'm going to write this in the following suggestive way. This is equal to 1 times f of x minus L derivative with respect to x f of x plus L squared over 2 derivative with respect to x squared times f of x minus L cubed over 6 derivative cubed with respect to x plus dot, dot, dot. Everybody good with that?

But this is a series that you should recognize, a particular Taylor series for a particular function. It's a Taylor expansion for the

AUDIENCE: Exponential.

PROFESSOR: Exponential. e to the minus L derivative with respect to x of f of x . Which is kind of awesome. So let's just check to make sure that this makes sense from dimensional grounds. So that's a derivative with respect to x as units of 1 over length. That's a length, so this is dimensionless, so we can exponentiate it.

Now you might look at me and say, look, this is silly. You've taken an operation like derivative and exponentiated it. What does that mean? And that is what it means?

[LAUGHTER]

OK? So we're going to do this all the time in quantum mechanics. We're going to do things like exponentiate operations. We'll talk about it in more detail, but we're always going to define it in this fashion as a formal power series. Questions?

AUDIENCE: Can you transform operators from one space to another?

PROFESSOR: Oh, you totally can. But we'll come back to that. We're going to talk about operators next time.

OK, so here's where we are. So from this what is a derivative with respect to x mean? What does a derivative with respect to x do? Well a derivative with respect to x is something that generates translations with respect to x through a Taylor expansion.

If we have L be arbitrarily small, right? L is arbitrarily small. What is the translation by an arbitrarily small amount of f of x ? Well, if L is arbitrarily small, we can drop all the higher order terms, and the change is just Ldx . So the derivative with respect to x is telling us about infinitesimal translations. Cool?

The derivative with respect to a position is something that tells you, or controls, or generates infinitesimal translations. And if you exponentiate it, you do it many, many, many times in a particular way, you get a macroscopic finite translation. Cool?

So this gives us three things. Translations in x are generated by derivative with respect to x . But through Noether's theorem translations, in x are associated to conservation of momentum. So you shouldn't be so shocked-- it's really not totally shocking-- that in quantum mechanics, where we're very interested in the action of things on functions, not just in positions, but on functions of position, it shouldn't be totally shocking that in quantum mechanics, the derivative with respect to x is related to the momentum in some particular way.

Similarly, translations in t are going to be generated by what operation? Derivative with respect to time. So derivative with respect to time from Noether's theorem is associated with conservation of energy. That seems plausible. Derivative with respect to, I don't know, an angle, a rotation. That's going to be associated with what? Angular momentum? But angular momentum around the axis for whom this is the angle, so I'll call that z for the moment. And we're going to see these pop up over and over again.

But here's the thing. We started out with these three principles today, and we've let ourselves to some sort of association between the momentum and the derivative like this. OK? And I've given you some reason to believe that this isn't totally insane. Translations are deeply connected with conservation of momentum. Transitional symmetry is deeply connected with conservation momentum. And an infinitesimal translation is nothing but a derivative with respect to position. Those are deeply linked concepts.

But I didn't derive anything. I gave you no derivation whatsoever of the relationship between d/dx and the momentum. Instead, I'm simply going to declare it. I'm going to declare that, in quantum mechanics-- you cannot stop me-- in quantum mechanics, p is represented by an operator, it's represented by the specific operator \hbar upon i derivative with respect to x . And this is a declaration. OK?

It is simply a fact. And when they say it's a fact, I mean two things by that. The first is it is a fact that, in quantum mechanics, momentum is represented by derivative with respect to x times \hbar upon i . Secondly, it is a fact that, if you take this

expression and you work with the rest of the postulates of quantum mechanics, including what's coming next lecture about operators and time evolution, you reproduce the physics of the real world. You reproduce it beautifully.

You reproduce it so well that no other models have even ever vaguely come close to the explanatory power of quantum mechanics. OK? It is a fact. It is not true in some epistemic sense. You can't sit back and say, ah a priori starting with the integers we derive that p is equal to-- no, it's a model.

But that's what physics does. Physics doesn't tell you what's true. Physics doesn't tell you what a priori did the world have to look like. Physics tells you this is a good model, and it works really well, and it fits the data. And to the degree that it doesn't fit the data, it's wrong. OK? This isn't something we derive. This is something we declare. We call it our model, and then we use it to calculate stuff, and we see if it fits the real world.

Out, please, please leave. Thank you.

[LAUGHTER]

I love MIT. I really do. So let me close off at this point with the following observation.

[LAUGHTER]

We live in a world governed by probabilities. There's a finite probability that, at any given moment, that two pirates might walk into a room, OK?

[LAUGHTER]

You just never know.

[APPLAUSE]

But those probabilities can be computed in quantum mechanics. And they're computed in the following ways. They're computed the following ways as we'll study in great detail. If I take a state, ψ of x , which is equal to e^{ikx} , this is a state

that has definite momentum $\hbar k$. Right? We claimed this. This was de Broglie and Davisson-Germer.

Note the following, take this operator and act on this wave function with this operator. What do you get? Well, we already know, because we constructed it to have this property. \hat{p} on $\psi(x)$ and I'm going to call this $\psi_k(x)$, because it has a definite k is equal to $\hbar k \psi_k(x)$.

A state with a definite momentum has the property that, when you hit it with the operation associated with momentum, you get back the same function times a constant, and that constant is exactly the momentum we ascribe to that plane wave. Is that cool? Yeah?

AUDIENCE: Question. Just with notation, what does the hat above the p [INAUDIBLE]?

PROFESSOR: Good. Excellent. So the hat above the P is to remind you that P is on a number. It's an operation. It's a rule for acting on functions. We'll talk about that in great detail next time.

But here's what I want to emphasize. This is a state which is equal to all others in the sense that it's a perfectly reasonable wave function, but it's more equal because it has a simple interpretation. Right? The probability that I measure the momentum to be $\hbar k$ is one, and the probability that I measure it to be anything else is 0, correct?

But I can always consider a state which is a superposition. Ψ is equal to α , let's just do $\frac{1}{\sqrt{2}} e^{ikx} + \frac{1}{\sqrt{2}} e^{-ikx}$.

Is this state a state with definite momentum? If I act on this state-- I'll call this ψ sub s -- if I act on this state with the momentum operator, do I get back this state times a constant? No. That's interesting. And so it seems to be that if we have a state with definite momentum and we act on it with momentum operator, we get back its momentum.

If we have a state that's a superposition of different momentum and we act on it with

a momentum operator, this gives us $\hbar k_1$, this gives us $\hbar k_2$. So it changes which superposition we're talking about. We don't get back our same state.

So the action of this operator on a state is going to tell us something about whether the state has definite value of the momentum. And these coefficients are going to turn out to contain all the information about the probability of the system. This is the probability when norm squared that will measure the system to have momentum k_1 . And this coefficient norm squared is going to tell us the probability that we have momentum k_2 .

So I think the current wave function is something like a superposition of $\frac{1}{10}\psi_{\text{pirates}}$ plus $\frac{1}{100}\psi_{\text{no pirates}}$. To normalize it properly $\psi_{\text{no pirates}}$. And I'll leave you with pondering this probability. See you guys next time.

[APPLAUSE]

CHRISTOPHER We've come for Prof. Allan Adams.

SMITH:

PROFESSOR: It is I.

CHRISTOPHER When in the chronicles of wasted time, I see descriptions of fairest rights, and I see
SMITH: lovely shows of lovely dames. And descriptions of ladies dead and lovely nights.
Then in the bosom of fair loves depths. Of eyes, of foot, of eye, of brow. I see the
antique pens do but express the beauty that you master now. So are all their
praises but prophecies of this, our time. All you prefiguring. But though they had but
diving eyes--

PROFESSOR: I was wrong about the probabilities.

[LAUGHTER]

CHRISTOPHER But though they had but diving eyes, they had not skill enough you're worth to sing.

SMITH: For we which now behold these present days have eyes to behold.

[LAUGHTER]

But not tongues to praise.

[APPLAUSE]

It's not over. You wait.

ARSHIA SURTI: Not marbled with gilded monuments of princes shall outlive this powerful rhyme. But you shall shine more bright in these contents that unswept stone besmear its sluttish tide. When wasteful war shall statues overturn and broils root out the work of masonry. Nor Mars his sword. Nor war's quick fire shall burn the living record of your memory. Gainst death and all oblivious enmity shall you pace forth. Your praise shall still find room, even in the eyes of all posterity. So no judgment arise till you yourself judgment arise. You live in this and dwell in lover's eyes.

[APPLAUSE]

CHRISTOPHER SMITH: Verily happy Valentine's day upon you. May your day be filled with love and poetry. Whatever state you're in, we will always love you.

[LAUGHTER]

[APPLAUSE]

Signed, Jack Florian, James [INAUDIBLE].

[LAUGHTER]

PROFESSOR: Thank you, sir. Thank you.

CHRISTOPHER SMITH: Now we go.

SMITH:

[APPLAUSE]