

An Introduction to Quantum Mechanics

February 5, 2007

1 Quantum Theory and Wave Particle Duality

What is quantum mechanics? The two-word moniker for the field is remarkably unenlightening as a definition. The quantum part seems to have something to do with “quantization” or rigid enumeration. And the mechanics part would indicate that it has something to do with the “mechanism” of motion of a physical system. While these elements may be the outcome of what quantum mechanics *does*, it really has nothing to do with what the field *is*.

I would define the field of quantum mechanics this way:

Quantum mechanics is a logical discipline aimed at forming consistent arguments based on *apparently* contradictory premises.

I italicize the adjective “apparently” because the contradictory premises referred to “really” aren’t contradictory (whatever really *really* means!) Usually the premises referred to are part of the **wave/particle (WP)paradox**.

Normally, we like to think of particles and waves as separate physical “objects.” Particles are localized and have mass. Waves are “spread out” spatially, have no mass and can “interfere” with one another to form diffraction patterns. At the turn of the last century, a series of major experiments showed that sometimes waves acted like particles, and visa versa. That’s where the “paradox” came in. Let’s look at two of these experiments. We start with the photoelectron experiment (explained by Einstein - it got him the nobel prize!)

If you shine a light on a surface, you can get electrons to leave the solid through the illuminated surface. This is called photoemission. The interesting thing about photoemission is that there is a critical light wavelength¹ at which the emission begins. Longer wavelengths produce no photoemission. Shorter wavelengths give more photoemission. If the emission process required some critical amount of energy incident, as you’d think “classically,” just upping the intensity of the longer wavelengths should get electrons out of the surface. But this is not true. Upping intensity doesn’t get electrons out.

¹or frequency, since $\lambda\nu = c$ where λ = wavelength, ν = frequency and c is the speed of light

Einstein “explained” this effect by saying that light has particle-like properties. That is, it was composed of the old, pre-Newtonian “corpuscles,” which Einstein (I think) renamed “photons.” Each photon had a certain amount of “oomph” (energy) which it could transfer to individual atoms to knock their electrons out. The energy of the photon was proportional to the frequency:

$$E = h\nu = \hbar\omega \quad (1)$$

To summarize, the picture Einstein painted was this. A “particle” of light (a photon) acts on a single atom, imparting an amount of energy, $\hbar\omega$ to it. The photons interacted with the electrons in the atom. If the photon energy exceeded the “binding energy” of the electron, the electron would be ejected from the atom and leave the solid through the illuminated surface. Longer wavelength light could interact with the solid in different ways, either passing through the solid (transparency) or heating the solid. So it seems that what we generally think of as an electromagnetic wave (light) has some particle properties.

The next of these “blockbuster” experiments was conducted by Davidson and Germer at Bell Laboratories. Quite by accident, they discovered that electrons incident on a crystal could, apparently, diffract (interfere with on another constructively or destructively). The experiment they performed was remarkably like the experiments Bragg and Von Laue conducted by shining an x-ray beam on a crystal solid surface (see fig.1).

In this figure, I show reflection of a collimated (parallel-ray) x-ray beam from two successive planes of crystal. For the two reflections to form a coherent, constructively interfering emerging wavefront, the total path-length difference $2\Delta l$ must be an integral number of wavelengths. this yields the following formula for a constructively interfering emerging wave:

$$n\lambda = 2d \sin(\theta) \quad (2)$$

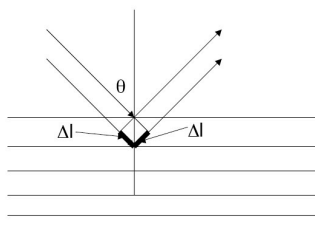


Figure 1: X-ray reflection from the top two crystallographic planes of a crystal.

Here, d is the interplanar spacing.

Think about the following experiment. Shine a collimated x-ray beam, filtered to provide 1 wavelength, on a solid. Think about the crystal as composed of planes, such as those shown in fig.1. Rotate the crystal about an axis normal to the page in fig. 1. There is a series of unique angles (as defined by integer n) where the crystal “flashes out” with reflected radiation. It doesn’t reflect at all angles, only those dictated by eq. 1.

What Davidson and Germer did was to show that when a beam of electrons was incident on a solid, the resulting “reflection” obeyed eq. 1 as well. But they did more than that. They verified the De Broglie equation:

$$p = \hbar k \quad (3)$$

as well. Here, p is the momentum of the “particle,” \hbar is the planck constant and k is a wave number given as:

$$k = \frac{2\pi}{\lambda} \quad (4)$$

Note: This equation mixes wave and particle properties. Momentum is usually associated with a particle of some mass and velocity. Of course, wavelength is a pure wave property.

To see how they did this, consider that a classical electron moving through field-free space has a purely kinetic energy, E . This energy is:

$$E = \frac{1}{2}mv^2 = \frac{1}{2m}m^2v^2 = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 4\pi^2}{2m\lambda^2} \quad (5)$$

This allows us to solve for λ :

$$\lambda = \frac{\sqrt{2mE}}{2\pi\hbar} \quad (6)$$

Substitution into eq.2 gives a new Bragg condition for electrons:

$$\frac{n\sqrt{2mE}}{2\pi\hbar} = 2d \sin(\theta) \quad (7)$$

It appears that the sine of the diffraction angle θ is proportional to the square root of energy. The electron diffraction experiment verified this. And so, in one fell swoop, Davidson and Germer showed the wave-like nature of the electron and verified the De Broglie equation. Needless to say, this won the whole crew (De Broglie, Davidson and Germer) their nobel prizes.

And so, from these discussions, it appears that wave-like and particle-like behaviors are both associated with (or, they are properties of) objects as they appear in the physical universe. These properties can exist simultaneously in a given object. How these properties are manifest in actual observation depends on the type of observation we make. From the above discussion, it seems that we can merge these two properties in a single entity and

get a consistent result (look at eq. 7, for example). What appeared to be a contradiction at first, really isn't a contradiction at all.

But there is still that nagging question, "does the electron actually 'interfere' with itself, or, with other electrons?" To eliminate contradiction here, more quantum mechanics (in the sense of my opening definition) is needed. We have to be more precise by what we mean by the term "interfere."

2 The Copenhagen Interpretation and the Schrodinger Equation

Let us accept as fact the principle that matter and light have both wave-like and particle-like properties existing simultaneously within them. We don't have any problem seeing how waves can interfere. If we switch our thinking over to the "particle" side, we find our difficulties appearing. How can particles "interfere?" The answer, on the particle side, is: they don't! To accommodate this position, Bohr and others (who were living in Copenhagen² at the time. They said that our observations, our senses, mainly register the presence or absence of a particle and that there was a probability associated with registering this presence or absence. Some how, this probability could be made to exhibit the "light/dark" intensity pattern we associate with wave interference.

For this to become concrete and useful, we had to develop a mathematics which would reflect this. The mathematics was suggested by the "complex" math that describes diffractive interference. That is, for example, we can describe an electromagnetic wave in terms of its **unmeasurable** amplitude, or in terms of a measurable **intensity**, the power incident on a detector surface. The intensity is just the complex square of the wave amplitude.

Consider expressing a wave in complex form (ignoring the time-dependent part for a moment). Usually, we write:

$$\Psi_1(x) = A_1 \exp(ikx) \quad (8)$$

As usual, i is the square root of -1. Next, consider a second wave:

$$\Psi_2(x) = A_2 \exp(ik(x + \Delta x)) \quad (9)$$

and we allow this second wave to interact with the first. The rule here is that the interaction is represented as a simple sum of amplitudes

$$\Psi(x) = \Psi_1(x) + \Psi_2(x) = A_1 \exp(ikx) + A_2 \exp(ik(x + \Delta x)) \quad (10)$$

If $A_1 = A_2 = A$ and $\Delta x = 0$, we have:

²If you are interested in seeing what Copenhagen was like for Bohr and his colleagues, you can go to see the play "Copenhagen," or read the book.

$$\Psi(x) = 2A \exp(ikx) \quad (11)$$

The rule established above is that intensity is the complex square of the amplitude:

$$I = \Psi^*(x)\Psi(x) \quad (12)$$

where the star over the first Ψ indicates complex conjugation. For this simple case, the waves sum together constructively and the positional dependence of the intensity “goes away, and we have:

$$I = 4A^2 \quad (13)$$

a constant. (The reader should verify this is the case). Thus, for waves adding in phase, interference is only constructive and no “diffraction effects” appear.

Now consider the case in which Δx is not zero. For simplicity, we still take the amplitudes to be equal. We now have:

$$\Psi(x) = A \exp(ikx)(1 + \exp(ik\Delta x)) \quad (14)$$

The intensity is the complex square of this term:

$$I(x) = \Psi^*(x)\Psi(x) = 2A^2(1 + 2\cos(k\Delta x)) \quad (15)$$

Again, the reader is encouraged to fill in missing steps in the “complex” math. Note that the over all intensity is constant (not position dependent). But the intensity fluctuates from zero (complete destructive interference) to $4A$ (complete constructive interference), depending on the value of Δx .

In optical diffraction, though, you say you’ve seen intensity variation patterns which had spatial dependence. For our current example, there was no spatial dependence on the phase factor Δx . To reconstruct this spatial dependence consider the “two-slit” set-up in the following picture.

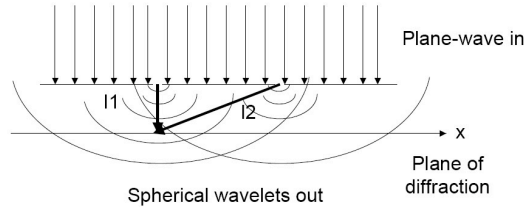


Figure 2: A two slit interference experiment demonstrating variation in intensity along x due to wavefront interference effects.

As you see in this figure, the phase difference between the spherical waves emitted from the point source “slits” is given by $\Delta x = l1 - l2$.

Problem 1: This analysis gives a light and dark pattern in the x-direction in the plane of diffraction. It is left as an exercise to the reader to work out what this pattern is.

Hopefully, this discussion has convinced you that the complex method of representing waves and the concept that amplitudes add, giving rise to interference phenomena, provides a mathematical framework for understanding diffraction effects. But, once again, the story isn’t finished. We must next make the connection between “interfering probabilities” and this complex arithmetic. We do this by writing a new equation, the Schrodinger equation, that provides the complex probability “amplitudes.” Squaring these amplitudes will give us the particle location probabilities and, at the same time, account for the **apparent** interference effects.

Our goal here is to provide a kind of “quantum analog” of Newton’s laws. To do this, we begin with Hamilton’s modification of Newton’s principles. Hamilton (and Lagrange before him) wanted to derive classical mechanics (the science of motion for particles and extended masses) from a “variational” principle. That is, these people wanted to define a “control function” for mechanical systems. When the system was in equilibrium, this control function would take on some smallest value. The path the system would take to equilibrium would be the “shortest” one between the initial and final points on a surface defined by the control function.

The control function Hamilton chose was pretty straight forward: it was just an expression for total energy of the system, E . We know from freshman physics that the total energy of a system is the sum of the kinetic (KE) and potential(V) energies of all the masses in the system:

$$E = KE + V \tag{16}$$

Sometimes, E is replaced by H and the equation is called “the system Hamiltonian.” Schrodinger set out to create a probability amplitude equation from this starting point. Also, as energy expressed this way is a generalized function of the system coordinates, he wanted to define an **operator** equation for the probability amplitudes. Remember, an operator is a mathematical object that “acts on a function,” just like a function is a mathematical object that acts on a number.

But Schrodinger wanted to do a bit more. He wanted to his equation to account for “wavelike” and “particle” properties of matter. So he started by assuming a wavelike solution to his equation:

$$\Psi(r, t) = A \exp[i(kx - \omega t)] \tag{17}$$

He also required that the wave-particle duality equations of Einstein and De Broglie ($E = \hbar\omega$ and $p = \hbar k$) would be embedded in his equation.

We proceed with these requirements in mind. Remember some examples of operators: differentiation (of any order) and integration. Consider the action of a differentiation operators on eq. 17

$$\frac{\partial \Psi(x, t)}{\partial x} = ik\Psi(x, t) \quad (18)$$

$$\frac{\partial \Psi(x, t)}{\partial t} = -i\omega\Psi(x, t) \quad (19)$$

Multiplying both sides of these two equations by \hbar , and, with minor rearrangement, we get

$$\frac{\hbar}{i} \frac{\partial \Psi(x, t)}{\partial x} = \hbar k\Psi(x, t) \quad (20)$$

$$-\frac{\hbar}{i} \frac{\partial \Psi(x, t)}{\partial t} = \hbar\omega\Psi(x, t) \quad (21)$$

Thus we can say that an “operator” notation for momentum, p , is (with reference to eq. 20):

$$p \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x} \quad (22)$$

and (with reference to eq.21):

$$E \rightarrow -\frac{\hbar}{i} \frac{\partial}{\partial t} \quad (23)$$

We can complete our operator equation for the probability amplitude by writing the semi-classical expression for the Hamiltonian:

$$E = \frac{p^2}{2m} + V \quad (24)$$

and using the operator equivalent terms for E and p , we get;

$$-\frac{\hbar}{i} \frac{\partial \Psi(x, t)}{\partial t} = V(x, t)\Psi(x, t) - \frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} \quad (25)$$

This is the famous Schrodinger equation. it is an equation for the probability **amplitudes** of particles in some potential field. Complex squaring the resulting probability amplitude will provide a measure of the probability of finding a particle at a given location in a given time. It looks very much like the classical heat flow equation (first order in time, second order in position). But the imaginary term in the left hand side makes it an entirely

different animal. The imaginary term is necessary to allow for solutions with a traveling wave property.

Note two things. The equation embeds wave and particle properties in a single equation.³ Also, all the equation does is provide us with a means to find location probabilities. It doesn't say anything about particle acceleration. All it gives is probabilities.

Thus, comments to the effect that the probability function “indicates that the particle spends more time here or there as it buzzes about in the potential space” are just flat out wrong! There's no mention of how particles arrive at the point of measurement. We just get a probability that they're there from the equation.⁴

We can also make the following “conclusions and observations” concerning our “researches” in this section. The first observation talks to the issue of why we have an equation for a probability amplitude, and not for a probability itself. Our goal in setting up the Schrodinger equation was to provide a mathematics that would neatly express wave interference effects. Think back on how Maxwell's E&M equations allow this to be done. The basic equations solve for field amplitudes. These sum vectorially in three dimensional space and lead to the light and dark patterns associated with diffraction. As we saw above, an individual wave converts to a spatially homogeneous energy deposition pattern on a plane of observation. This energy is expressed as a non-negative scalar.

Once a light wave deposits its energy (through absorption, say, on a solid surface) phase information is lost and diffraction is not possible. We say that amplitudes interfere, intensities sum. The essence of this statement is at the core of the Copenhagen interpretation. Wave functions interfere, probabilities imply sum. The phase information creates the true interference. At the time of measurement, this information is lost and only intensity (or probability) remains.

So what exactly is “the Copenhagen Interpretation?” After all, Schrodinger invented the equation. The physicist Max Born was actually responsible for the probabilistic interpretation of the wavefunction. Sometimes the uncertainty relationships:

$$\Delta x \Delta p \approx h \tag{26}$$

$$\Delta t \Delta E \approx h \tag{27}$$

(where h is the Planck constant) gets lumped into the mix. And so does the “Correspondence Principle,” which states that classical behavior can be recovered in some “limiting” sense (i.e, as the quantum numbers get really large, or as h gets small.)

³Sometimes, this “wedding” of wave and particle properties is known as “complementarity.”

⁴Sometimes the apparent “collapse” of an “extended” wave function to a point particle at the time of measurement is called “the collapse of the wavefunction.” This, too, is not correct. The probability function (and its associated amplitude) are still relevant ways to help define a system. As long as the system exists, the wavefunction doesn't “collapse.” . There is a valid concept of wavefunction collapse. But this relates to the selection of a single quantum mechanical state from a “superposition” of states. This has to do with “quantum entanglement” and the “Schrodinger's Cat Paradox.” All of this is very interesting, but beyond the scope of our current discussion.

I guess the main ingredient of a “Copenhagen Interpretation” is the realization that the methods of quantum mechanics do not say anything about the “underlying reality” of the physical system. Quantum mechanics merely says something about how likely we are to make a certain observation, not how that observation came about. Also, we cannot say with any real authority that an electron will be “here” on July 5, 2080 at 10:51:13:34...PM. This drove Einstein nuts (well, nuttier than he was before his encounters with Bohr, Schrodinger, Heisenberg and Born). He spend a large portion of his career pushing his “god doesn’t play dice with the universe” emotion forward.

While Einstein’s criticism was very constructive and has led to many interesting modern concepts (like “quantum entanglement,” again beyond our current scope), I always thought his emotional reaction was a bit too strong. After all, Bohr’s emphasis in the interpretation was on **observable** quantities. What counted was the observation. He showed how, by merely finding a way to derive the likelihood of an observation, we’ve gotten what we want out of the physics anyway.

3 Some Applications

Now let us turn our attention to how we can get useful information out of this new quantum formalism.

3.1 Expectation Values

One of the first things we can do is use the theory as developed to calculate some system parameters. These parameters might answer questions like “where am I most likely to find a particle?” Or, “what is the most likely energy you’d measure for a particle after a group of repeated measurements?” In both of these cases, we are asking to find something called the “expectation” value of a parameter.

From elementary probability theory, we are tempted to write the most likely value as the average value you obtain after a “large” number of parameter measurements made on the same system under the same conditions. If $p(x)$ is the (normalized) probability of finding a value of x for a measurement (in the range from x to $x + dx$), we’d write the “expected” value of x as:

$$\langle x \rangle = \int_0^{\infty} xp(x)dx \quad (28)$$

where the wedge brackets mean “expected value” (of x , in this case). But that’s not how it works in quantum mechanics.

Consider finding the average momentum, $\hbar k$ for a free electron in a very large 1D box of side L . If you used our operator formalism, and take the electron to be represented by a “wavelike” wavefunction:

$$\Psi(x) = \sqrt{\frac{1}{L}} \exp[i(kx - \omega t)] \quad (29)$$

we'd have:

$$\langle p \rangle = \int_0^L \frac{\hbar}{i} \frac{\partial}{\partial k} \Psi^*(x) \Psi(x) dx = 0 \quad (30)$$

We know this isn't true. We get around the problem by writing the expectation value as:

$$\langle p \rangle = \int_0^L \Psi^*(x) \frac{\hbar}{i} \frac{\partial}{\partial k} \Psi(x) dx = \hbar k \quad (31)$$

which is what we'd "expect" as $\hbar k$ is the De Broglie expression for p .

The reason this works out is that we construct our operators so that they return real numbers when they operate on wavefunctions. Once they've done their job, they "disappear" and were left with an integral whose kernel is a real number times the complex conjugate of the wave function times the wave function itself. Clearly, the resulting integral is a real number - an observable.

Before we move on, let's question our methods in greater depth.

Problem 2:

- a. Note that the exponential in eq. 29 has a $1/\sqrt{L}$ pre-factor. Why?
- b. Note that the limits of the integration in eq. 30 are between 0 and ∞ . Why?

Now let's deal with some other expectation issues. What's the expected position of a particle with a wave function expressed like eq. 29?

$$\langle x \rangle = \int_0^L \sqrt{\frac{1}{L}} \exp[-i(kx - \omega t)] x \sqrt{\frac{1}{L}} \exp[i(kx - \omega t)] dx = \frac{1}{L} \frac{x^2}{2} \Big|_0^L = \frac{L}{2} \quad (32)$$

Another interesting expectation value is the position of the particle in the box with infinitely high walls, like we worked in class - the artificial atom problem. The wavefunction we derived for an electron in the box is:

$$\Psi(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \quad (33)$$

and:

$$\langle x \rangle = \int_0^L \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) x \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) dx = \frac{2}{L} \int_0^L x \sin^2\left(\frac{n\pi x}{L}\right) dx \quad (34)$$

The last integral on the right works out to $L/2$, just like the traveling wave.