Chapter 40

Quantum Mechanics

That light had both a particle and a wave nature became apparent with Einstein's explanation of the photoelectric effect in 1905. One might expect that such a discovery would lead to a flood of publications speculating on how light could behave both as a particle and a wave. But no such response occurred. The particle wave nature was not looked at seriously for another 18 years, when de Broglie proposed that the particle wave nature of the electron was responsible for the quantized energy levels in hydrogen. Even then there was great reluctance to accept de Broglie's proposal as a satisfactory thesis topic.

Why the reluctance? Why did it take so long to deal with the particle-wave nature, first of photons then of electrons? What conceptual problems do we encounter when something behaves both as a particle and as a wave? How are these problems handled? That is the subject of this chapter.

TWO SLIT EXPERIMENT

Of all the experiments in physics, it is perhaps the 2 slit experiment that most clearly, most starkly, brings out the problems encountered with the particle-wave nature of matter. For this reason we will use the 2 slit experiment as the basis for much of the discussion of this chapter.

Let us begin with a review of the 2 slit experiment for water and light waves. Figure (1) shows the wave pattern that results when water waves emerge from 2 slits. The lines of nodes are the lines along which the waves from one slit just cancel the waves coming from the other. Figure (2) shows our analysis of the 2 slit pattern. The path length difference to the first minimum must be half a wavelength $\lambda/2$. This gives us the two similar triangles shown in Figure (2). If y_{min} is much less than D, which it is for most 2 slit experiments, then the hypotenuse of the big triangle is approximately D and equating corresponding sides of the similar triangles gives us the familiar relationship

$$\frac{\lambda/2}{d} = \frac{y_{min}}{D}$$

$$\boxed{\lambda = \frac{2y_{min}d}{D}} \qquad (1)$$

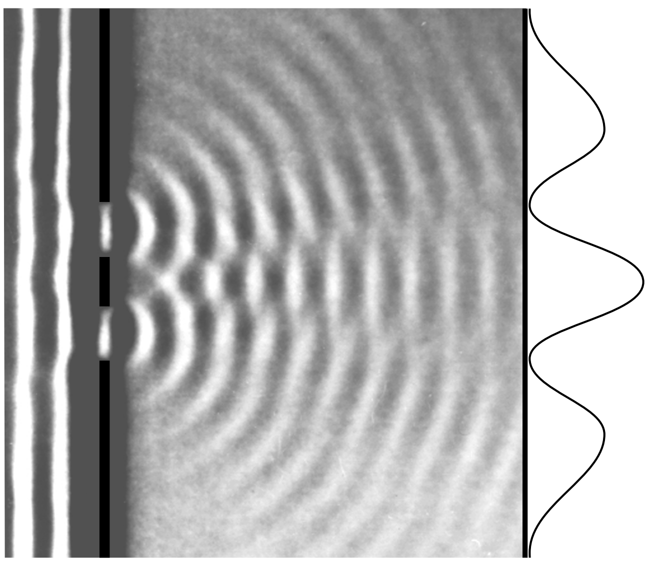


Figure 1
Water waves emerging from two slits.

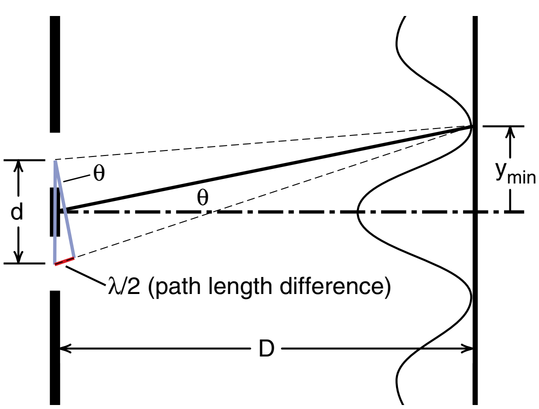


Figure 2
Analysis of the two slit pattern. We get a minimum when the path length difference is half a wavelength.

Figure (3a) is the pattern we get on a screen if we shine a laser beam through 2 slits. To prove that the dark bands are where the light from one slit cancels the light from the other, we have in Figure (3b) moved a razor blade in front of one of the slits. We see that the dark bands disappear, and we are left with a one slit pattern. The dark bands disappear because there is no longer any cancellation of the waves from the 2 slits.

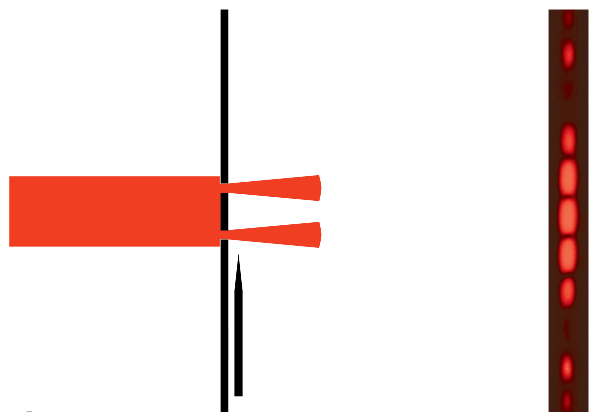


Figure 3a
Two slit interference pattern for light. The closely spaced dark bands are where the light from one slit cancels the light from the other.

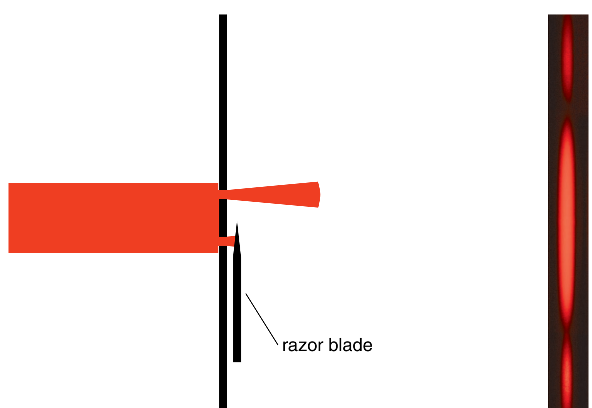


Figure 3b
Move a razor blade in front of one of the slits, and the closely spaced dark bands disappear. There is no more cancellation.

In 1961, Claus Jönsson did the 2 slit experiment using electrons instead of light, with the results shown in Figure (4). Assuming that the electron wavelength is given by the de Broglie formula $p = h/\lambda$, the dark bands are located where one would expect waves from the 2 slits to cancel. The 2 slit experiment gives the same result for light and electron waves.

The Two Slit Experiment from a Particle Point of View

In Figure (3a), the laser interference patterns were recorded on a photographic film. The pattern is recorded when individual photons of the laser light strike individual silver halide crystals in the film, producing a dark spot where the photon landed. Where the image shows up white in the positive print, many photons have landed close together exposing many crystal grains.

In a more modern version of the experiment one could use an array of photo detectors to count the number of photons landing in each small element of the array. The number of counts per second in each detector could then be sent to a computer and the image reconstructed on the computer screen. The result would look essentially the same as the photograph in Figure (3a).

The point is that the image of the two slit wave pattern for light is obtained by counting particles, not by measuring some kind of a wave height. When we look at the two slit experiment from the point of view of counting particles, the experiment takes on a new perspective.

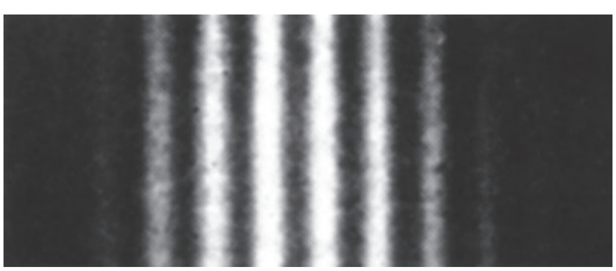


Figure 4
Two slit experiment using electrons. (By C. Jönsson)

Imagine yourself shrunk down in size so that you could stand in front of a small section of the photographic screen in Figure (3a). Small enough that you want to avoid being hit by one of the photons on the laser beam. As you stand at the screen and look back at the slits, you see photons being sprayed out of both slits as if two machine guns were firing bullets at you, but you discover that there is a safe place to stand. There are these dark bands where the particles fired from one slit cancel the particles coming from the other.

Then one of the slits is closed, there is no more cancellation, the dark bands disappear as seen in Figure (3b). There is no safe place to stand when particles are being fired at you from only one slit. It is hard to imagine in our large scale world how it would be safe to have two machine guns firing bullets at you, but be lethal if only one is firing. It is hard to visualize how machine gun bullets could cancel each other. But the particle wave nature of light seems to require us to do so. No wonder the particle nature of light remained an enigma for nearly 20 years.

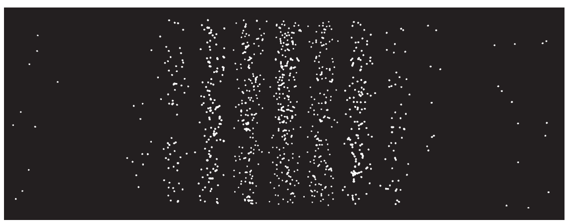
Two Slit Experiment—One Particle at a Time

You might object to our discussion of the problems involved in interpreting the two slit experiment. After all, Figure (1) shows water waves going through two slits and producing an interference pattern. The waves from one slit cancel the waves from the other at the lines of nodes. Yet water consists of particles—water molecules. If we can get a two slit pattern for water molecules, what is the big deal about getting a two slit pattern for photons ? Couldn't the photons somehow interact with each other the way water molecules do, and produce an interference pattern?

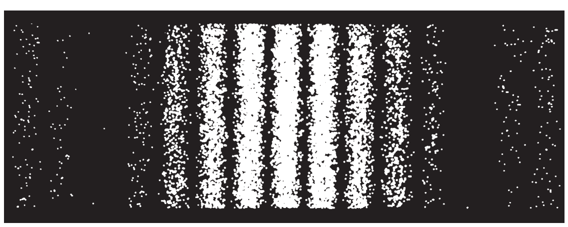
Photons do not interact with each other the way water molecules do. Two laser beams can cross each other with no detectable interaction, while two streams of water will splash off of each other. But one still might suspect that the cancellation in the two slit experiment for light is caused by some kind of interaction between the photons. This is even more likely in the case of electrons, which are strongly interacting charged particles.

a) 10 dots

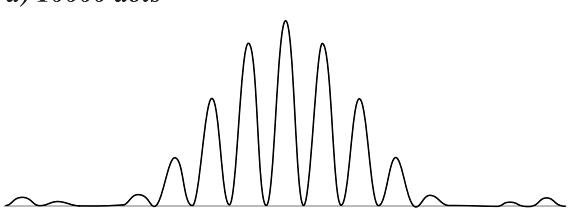
b) 100 dots



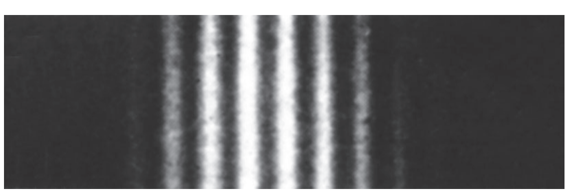
c) 1000 dots



d) 10000 dots



Predicted pattern



Experimental results by C. Jönsson

Figure 5
Computer simulation of the 2 slit electron diffraction experiment, as if the electrons had landed one at a time.

In an earlier text, we discussed the possibility of an experiment in which electrons would be sent through a two slit array, one electron at a time. The idea was to eliminate any possibility that the electrons could produce the two slit pattern by bouncing into each other or interacting in any way. Since the experiment had not yet been done, we drew a sketch of what the results should look like. That sketch now appears in a number of introductory physics texts.

When he saw the sketch, Lawrence Campbell of the Los Alamos Scientific Laboratories did a computer simulation of the experiment. We will first discuss Campbell's simulation, and then compare the simulation with the results of the actual experiment which was performed in 1991.

It is not too hard to guess some of the results of sending electrons through two slits, one at a time. After the first electron goes through you end up with one dot on the screen showing where the electron hit. The single dot is not a wave pattern. After two electrons, two dots; you cannot make much of a wave pattern out of two dots.

If, after many thousands of electrons have hit the screen, you end up with a two slit pattern like that shown in Figure (4), that means that none of the electrons land where there will eventually be a dark band. You know where the first dot, and the second dot cannot be located. Although two dots do not suggest a wave pattern, some aspects of the wave have already imposed themselves by preventing the dots from being located in a dark band.

To get a better idea of what is happening, let us look at Campbell's simulation in Figure (5). In (5a), and (5b) we see 10 dots and 100 dots respectively. In neither is there an apparent wave pattern, both look like a fairly random scatter of dots. But by the time there are 1000 dots seen in (5c), a fairly distinctive interference pattern is emerging. With 10,000 dots of (5d), we see a fairly close resemblance between Campbell's simulation and Jönsson's experimental results. Figure (5e) shows the wave pattern used for the computer simulation.

Although the early images in Figure (5) show nearly random patterns, there must be some order. Not only do the electrons not land where there will be a dark band, but they must also accumulate in greater numbers where the brightest bands will eventually be. If this were a roulette type of game in Las Vegas, you should put your money on the center of the brightest band as being the location most likely to be hit by the next electron.

Campbell's simulation was done as follows. Each point on the screen was assigned a probability. The probability was set to zero at the dark bands and to the greatest value in the brightest band. Where each electron landed was randomly chosen, but a randomness governed by the assigned probability.

How to assign a probability to a random event is illustrated by a roulette wheel. On the wheel, there are 100 slots, of which 49 are red, 49 black and 2 green. Thus where the ball lands, although random, has a 49% chance of being on red, 49% on black, 2% on green, and 0% on blue, there being no blue slots.

In the two slit simulation, the probability of the electron landing at some point was proportional to the intensity of the two slit wave pattern at that point. Where the wave was most intense, the electron is most likely to land. Initially the pattern looks random because the electrons can land with roughly equal probability in any of the bright bands. But after many thousands of electrons have landed, you see the details of the two slit wave pattern. The dim bands are dimmer than the bright ones because there was a lower probability that the electron could land there.

Figure (6) shows the two slit experiment performed in 1991 by Akira Tonomura and colleagues. The experiment involved a novel use of a superconductor for the two slits, and the incident beam contained so few electrons per second that no more than one electron was between the slits and the screen at any one time. The screen consisted of an array of electron detectors which recorded the time of arrival of each electron in each detector. From this data the researchers could reconstruct the electron patterns after 10 electrons (6a), 100 electrons (6b), 3000 electrons (6c), 20,000 electrons (6d) and finally after 70,000 electrons in Figure (6e). Just as in Campbell's simulation, the initially random looking patterns emerge into the full two slit pattern when enough electrons have hit the detectors.

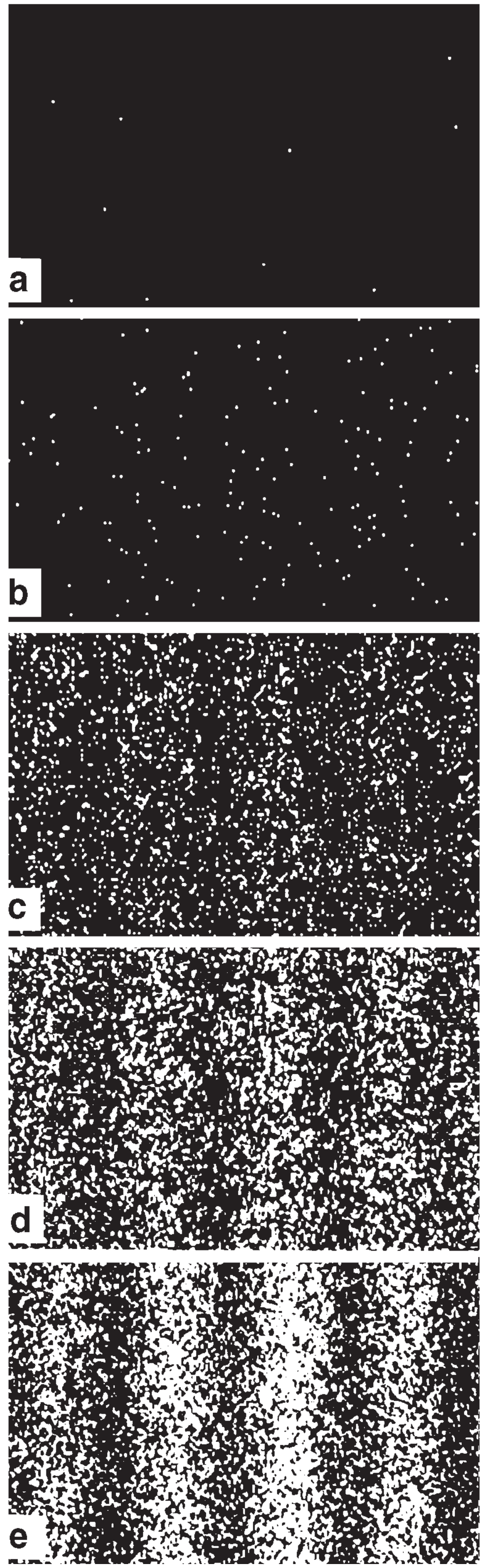


Figure 6
Experiment in which the 2 slit electron interference pattern is built up one electron at a time. (A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki, American Journal of Physics, Feb. 1989. See also Physics Today, April 1990, Page 22.)

Born's Interpretation of the Particle Wave

In 1926, while calculating the scattering of electron waves, Max Born discovered an interpretation of the electron wave that we still use today. In Born's picture, the electron is actually a particle, but it is the electron wave that governs the behavior of the particle. The electron wave is a probability wave governing the probability of where you will find the electron.

To apply Born's interpretation to the two slit electron experiment, we do what Campbell did in the simulation of Figure (5). We first calculate what the wave pattern at the screen would be for a wave passing through the two slits. It is the two slit interference pattern we have seen for water waves, light waves and electron waves. ***We then interpret the intensity of the pattern at some point on the screen as being proportional to the probability that the electron will land at that point.*** We cannot predict where any given electron will actually land, any more than we can predict where the ball will end up on the roulette wheel. But we can predict what the pattern will look like after many electrons have landed. If we repeat the experiment, the electrons will not land in the same places, but eventually the same two slit pattern will result.

Exercise 1

Figure (36-16) reproduced here, shows the diffraction pattern produced when a beam of electrons is scattered by the atoms of a graphite crystal. Explain what you would expect to see if the electrons went through the graphite crystal one at a time and you could watch the pattern build up on the screen. Could you market this apparatus in Las Vegas, and if so, how would you use it?

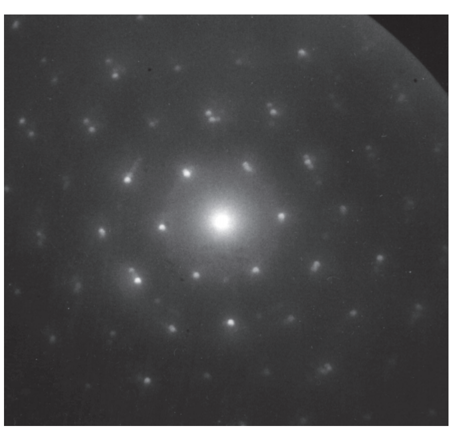


Figure 36-16 *Diffraction pattern produced by electrons passing through a graphite crystal.*

Photon Waves

Both electrons and photons have a particle-wave nature related by the de Broglie formula $p = h/\lambda$, and both produce a two slit interference pattern. Thus one would expect that the same probability interpretation should apply to electron waves and light waves.

We have seen, however, that a light wave, according to Maxwell's equations, consists of a wave of electric and magnetic fields $\vec{E}$ and $\vec{B}$. These are vector fields that at each point in space have both a magnitude and a direction. Since probabilities do not point anywhere, we cannot directly equate $\vec{E}$ and $\vec{B}$ to some kind of probability.

To see how to interpret the wave nature of a photon, let us first consider something like a radio wave or a laser beam that contains many billions of photons. In our discussion of capacitors in Chapter 27, we saw that the energy density in a classical electric field was given by

$$\left.\begin{matrix}\text{Energy}\\ \text{density}\end{matrix}\right\} = \frac{\varepsilon_0 E^2}{2} \qquad \textit{energy density in an electric field} \qquad (27\text{-}36)$$

where $E^2 = \vec{E}\cdot\vec{E}$. In an electromagnetic wave there are equal amounts of energy in the electric and the magnetic fields. Thus the energy density in a classical electromagnetic field is twice as large as that given y Equation 27-36, and we have

$$E\,\frac{\text{joules}}{\text{meter}^3} = \varepsilon_0 E^2 \qquad \textit{energy density in an electricomagnetic wave}$$

If we now picture the electromagnetic wave as consisting of photons whose energy is given by Einstein's photoelectric formula

$$E_{photon} = hf\,\frac{\text{joules}}{\text{photon}}$$

then the density of photons in the wave is given by

$$n = \frac{\varepsilon_0 E^2\ \text{joules}/\text{meter}^3}{hf\ \ \text{joules}/\text{photon}}$$

$$n = \frac{\varepsilon_0 E^2}{hf}\,\frac{\text{photons}}{\text{meter}^3} \qquad \textit{density of photons in an electromagnetic wave of frequency f} \qquad (1)$$

where f is the frequency of the wave.

In Exercise 2, we have you estimate the density of photons one kilometer from the antenna of the student AM radio station at Dartmouth College. The answer turns out to be around .25 billion photons/cc, so many photons that it would be hard to detect them individually.

Exercise 2

To estimate the density of photons in a radio wave, we can, instead of calculating $\vec{E}$ for the wave, simply use the fact that we know the power radiated by the station. As an example, suppose that we are one kilometer away from a 1000 watt radio station whose frequency is $1.4 \times 10^6 Hz$. A 1000 watt station radiates 1000 joules of energy per second or 10^{-6} joules in a nanosecond. In one nanosecond the radiated wave moves out one foot or about 1/3 of a meter. If we ignore spatial distortions of the wave, like reflections from the ground, etc., then we can picture this 10^{-6} joules of energy as being located in a spherical shell 1/3 of a meter thick, expanding out from the antenna.

(a) What is the total volume of a spherical shell 1/3 of a meter thick and 1 kilometer in radius?

(b) What is the average density of energy, in joules/m^3 of the radio wave 1 kilometer from the antenna

(c) What is the energy, in joules, of one photon of frequency $1.4 \times 10^6 Hz$?

(d) What is the average density of photons in the radio wave 1 kilometer from the station? Give the answer first in photons/m^3 and then photons per cubic centimeter. (The answer should be about .25 billion photons/cm^3.)

Now imagine that instead of being one kilometer from the radio station, you were a million kilometers away. Since the volume of a spherical shell 1/3 of a meter thick increases as r^2, (the volume being $(1/3) \times 4\pi r^2$) the density of photons would decrease as $1/r^2$. Thus if you were 10^6 times as far away, the density of photons would be 10^{-12} times smaller. At one million kilometers, the average density of photons in the radio wave would be

$$\left.\begin{array}{l}\text{number of photons}\\ \text{per cubic centimeter}\\ \text{at 1 million kilometers}\end{array}\right\} = \frac{\text{number at 1km}}{10^{12}} = \frac{.25 \times 10^9}{10^{12}} = .00025\,\frac{\text{photons}}{\text{cm}^3} \qquad (2)$$

In the classical picture of Maxwell's equations, the radio wave has a continuous electric and magnetic field even out at 1 million kilometers. You could calculate the value of $\vec{E}$ and $\vec{B}$ out at this distance, and the result would be sinusoidally oscillating fields whose structure is that shown back in Figure (32-23). But if you went out there and tried to observe something, all you would find is a few photons, on the order of .25 per liter (about one per gallon of space). If you look in 1 cubic centimeter of space, chances are you would not find a photon.

So how do you use Maxwell's equations to predict the results of an experiment to detect photons a million kilometers from the antenna? First you use Maxwell's equation to calculate $\vec{E}$ at the point of interest, then evaluate the quantity $(\varepsilon_0 \vec{E}\cdot\vec{E}/hf)$, and finally interpret the result as the probability of finding a photon in the region of interest. If, for example, we were looking in a volume of one cubic centimeter, the probability of finding a photon there would be about .00025 or .025%.

This is an explicit prescription for turning Maxwell's theory of electromagnetic radiation into a probability wave for photons. If the wave is intense, as it was close to the antenna, then $(\varepsilon_0 \vec{E}\cdot\vec{E}/hf)$ represents the density of photons. If the wave is very faint, then $(\varepsilon_0 \vec{E}\cdot\vec{E}/hf)$ becomes the probability of finding a photon in a certain volume of space.

Reflection and Fluorescence

An interesting example of the probability interpretation of light waves is provided by the phenomena of reflection and of fluorescence.

When a light beam is reflected from a metal surface, the angle of reflection, labeled θ_r in Figure (7a) is equal to the angle of incidence θ_i. The reason for this is seen in Figure (7b). The incident light wave is scattered by many atoms in the metal surface. The scattered waves add up to produce the reflected wave as shown in Figure (7b). Any individual photon in the incident wave must have an equal probability of being scattered by all of these atoms in order that the scattered probability waves add up to the reflected wave shown in (7b).

When you have a fluorescent material, you see a rather uniform eerie glow rather than a reflected wave. The light comes out in all directions as in Figure (8a).

The wavelength of the light from a fluorescent material is not the same wavelength as the incident light. What happens is that a photon in the incident beam strikes and excites an individual atom in the material. The excited atom then drops back down to the ground state radiating two or more photons to get rid of the excitation energy. (Ultraviolet light is often used in the incident beam, and we see the lower energy visible photons radiated from the fluorescing material.)

The reason that fluorescent light emerges in many directions rather than in a reflected beam is that an individual photon in the incident beam is absorbed by and excites one atom in the fluorescent material. There is no probability that it has struck any of the other atoms. The fluorescent light is then radiated as a circular wave from that atom, and the emerging photon has a more or less equal probability of coming out in all directions above the material.

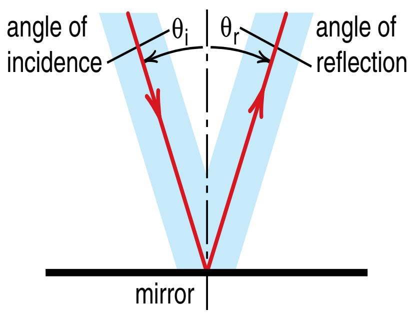


Figure 7a
When a light wave strikes a mirror, the angle of incidence equals the angle of reflection.

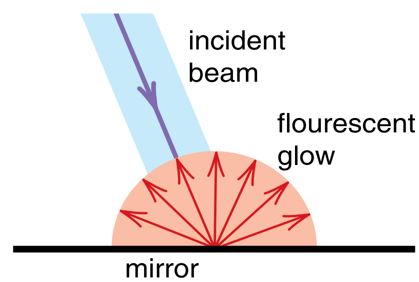


Figure 8a
When a beam of light strikes a fluorescent material, we see an eerie glow rather than a normal reflected light.

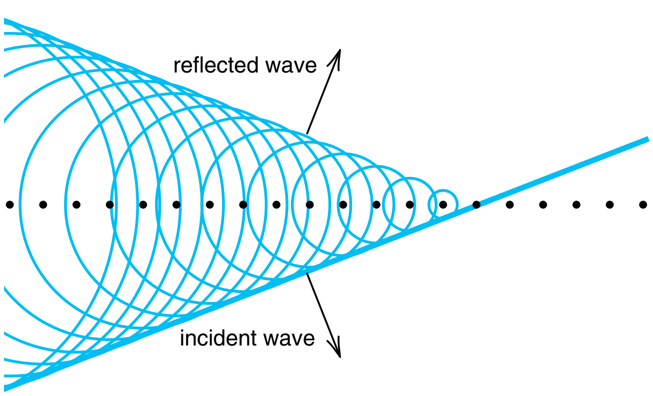


Figure 7b
The reflected wave results from the scattering of the incident wave by many atoms. If the incident wave contains a single photon, that photon must have an equal probability of being scattered by many atoms in order to emerge in the reflected wave.

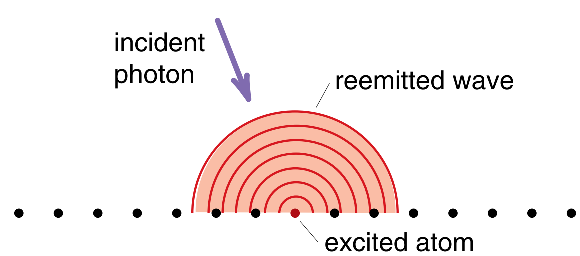


Figure 8b
Fluorescence occurs when an individual atom is excited and radiates its extra energy as two distinct photons. Since there is no chance that the radiation came from other atoms, the radiated wave emerges only from the excited atom.

A Closer Look at the Two Slit Experiment

While the probability interpretation of electron and photon waves provides a reasonable explanation of some phenomena, the interpretation is not without problems. To illustrate what these problems are, consider the following thought experiment.

Imagine that we have a large box with two slits at one end and a photographic film at the other, as shown in Figure (9). Far from the slits is an electron gun that produces a weak beam of electrons, so weak that on the average only one electron per hour passes through the slits and strikes the film. For simplicity we will assume that the electrons go through the slits on the hour, there being the 9:00 AM electron, the 10:00 AM electron, etc.

The electron gun is one of the simple electron guns we discussed back in Chapter 28. The beam is so spread out that there is no way it can be aimed at one slit or the other. Our beam covers both slits, meaning that each of the electrons has an equal chance of going through the top or bottom slit.

We will take the probability interpretation of the electron wave seriously. If the electron has an equal probability of passing through either slit, then an equally intense probability wave must emerge from both slits. When the probability waves get to the photographic film, there will be bands along which waves from one slit cancel waves from the other, and we should eventually build up a two slit interference pattern on the film.

Suppose that on our first run of the thought experiment, we do build up a two slit pattern after many hours and many electrons have hit the film.

We will now repeat the experiment with a new twist. We ask for a volunteer to go inside the box, look at the slits, and see which one each electron went through. John volunteers, and we give him a sheet of paper to write down the results. To make the job easier, we tell him to just look at the bottom slit on the hour to see if the electron went through that slit. If for example he sees an electron come out of the bottom slit at 9:00 AM, then the 9:00 AM electron went through the bottom slit. If he saw no electron at 10:00 AM, then the 10:00 AM electron must have gone through the upper slit.

If John does his job carefully, what kind of a pattern should build up on the film after many electrons have gone through? If the 9:00 AM electron was seen to pass through the bottom slit, then there is no probability that it went through the top slit. As a result, a probability wave can emerge only from the bottom slit, and there can be no cancellation of probability waves at the photographic film. Since the 10:00 AM electron did not go through the bottom slit, the probability wave must have emerged only from the top slit and there again can be no cancellation of waves at the photographic film.

If John correctly determines which slit each electron went through, there can be no cancellation of waves from the two slits, and we have to end up with a one slit pattern on the film. Just the knowledge of which slit each electron went through has to change the two slit pattern into a one slit pattern. With Born's probability interpretation of electron waves, just the ***knowledge*** of which slit the electrons go through ***changes the result of the experiment***. Does this really happen, or have we entered the realm of metaphysics?

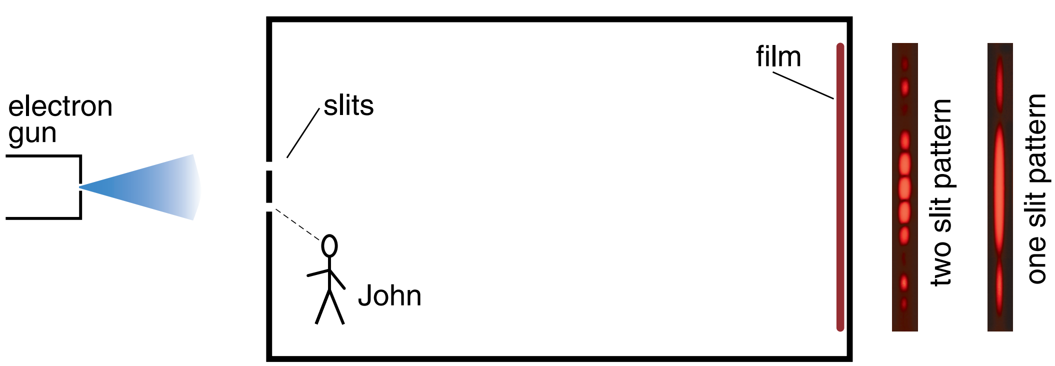


Figure 9
In this thought experiment, we consider the possibility that someone is looking at the two slits to see which slit each electron comes through.

Let us return to our thought experiment. John has been in the box for a long time now, so that a number of electrons have hit the film. We take the film out, develop it, and clearly see a two slit interference pattern emerging. There are the dark bands along which waves from one slit cancel the waves from the other slit.

Then we go over to the door on the side of the box, open it and let John out, asking to see his results. We look at his sheet of paper and nothing is written on it. "What were you doing all of that time," we ask. "What do you mean, what was I doing? How could I do anything? You were so careful sealing up the box from outside disturbances that it was dark inside. I couldn't see a thing and just had to wait until you opened the door. Not much of a fun experiment."

"Next time," John said, "give me a flashlight so I can see the electrons coming through the slits. Then I can fill out your sheet of paper."

"Better be careful," Jill interrupts, "about what kind of a flashlight you give John. A flashlight produces a beam of photons, and John can only see a passing electron if one of the flashlight's photons bounces off the electron."

"Remember that the energy of a photon is proportional to its frequency. If the photons from John's flashlight have too high a frequency, the photon hitting the passing electron will change the motion of the electron and mess up the two slit pattern. Give John a flashlight that produces low frequency, low energy photons, so he won't mess up the experiment."

"But," Bill responds, "a low frequency photon is a long wavelength photon. Remember that demonstration where waves were scattered from a tiny object? The scattered waves were circular, and contained no information about the shape of the object (Figure 36-1). You can't use waves to study details that are much smaller than the wavelength of the wave. That is why optical microscopes can't be used to study viruses that are smaller than a wavelength of visible light."

"If John's flashlight," Bill continues, "produced photons whose wavelength was longer than the distance between the two slits, then even if he hit the electron with one of the photons in the wave, John could not tell which slit the electron came through."

"Let us do some calculations," the professor says. "The most delicate way we can mess up the experiment is to hit an electron sideways, changing the electron's direction of motion so that if it were heading toward a maxima, it will instead land in a minima, filling up the dark bands and making the pattern look like a one slit pattern. Here is a diagram for the situation (Figure 10)."

"In the top sketch (10a), John is shining his flashlight at an electron that has just gone through the slit and is heading toward the central maximum. In the middle sketch (10b) the photon has knocked the electron sideways, so that it is now headed toward the first minimum in the diffraction pattern. Let us assume that all the photon's momentum $\vec{p}_{photon}$ has been transferred to the electron, so that the electron's new momentum is now

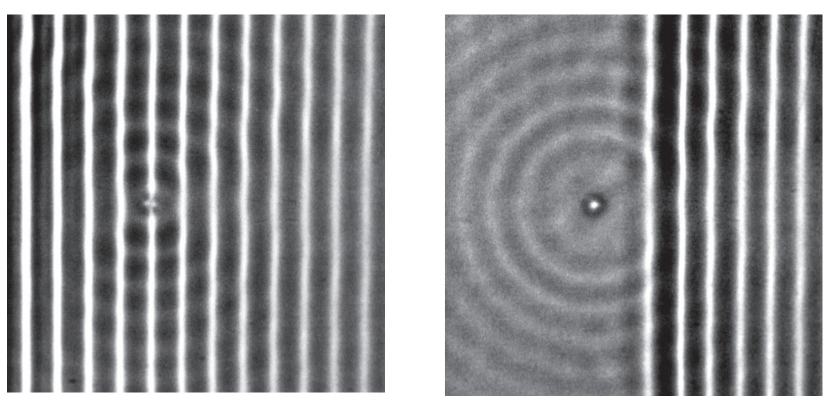


Figure 36-1 (reproduced)
If an object is smaller than a wavelength, the scattered waves are circular and do not contain information about the shape of the object.

Incident and scattered wave ***After incident wave has passed***

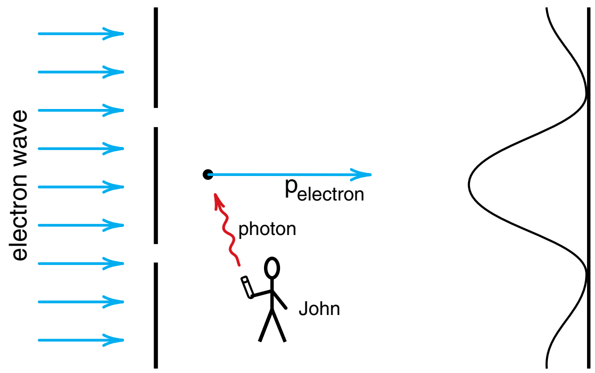


Figure 10a
In order to see the electron, John uses a flashlight, and strikes the electron with a photon.

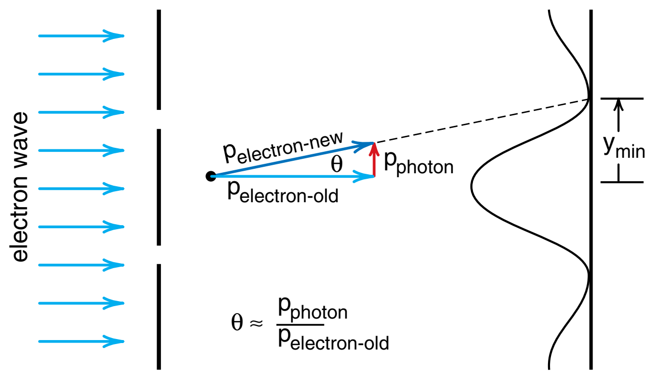


Figure 10b
Assume the photon's momentum has been absorbed by the electron. This could deflect the electron's path by an angle θ.

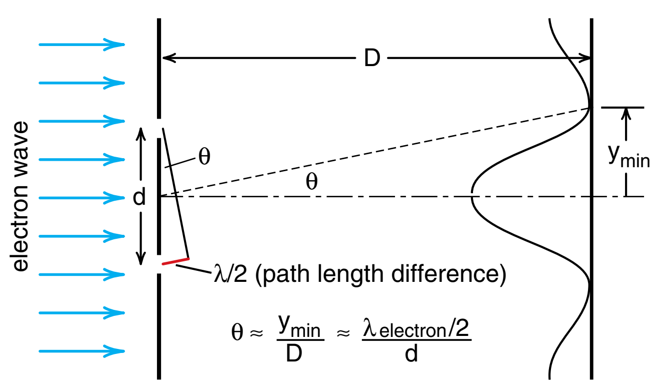


Figure 10c
Analysis of the two slit pattern. The angle to the first minimum is determined by using similar triangles. If the angle θ is small, then sin θ ≈ θ.

$$\vec{p}_{electron-new} = \vec{p}_{electron-old} + \vec{p}_{photon} \tag{3}$$

The angle θ by which the electron is deflected is approximately given by

$$\theta \approx \frac{p_{photon}}{p_{electron}} = \frac{h/\lambda_{photon}}{h/\lambda_{electron}}$$

$$\theta \approx \frac{\lambda_{electron}}{\lambda_{photon}} \tag{4}$$

where we used the de Broglie formula for the photon and electron momenta."

"In the bottom sketch we have the usual analysis of a two slit pattern. If the angle θ to the first minimum is small, which it usually is for a two slit experiment, then by similar triangles we have

$$\theta \approx \frac{y_{min}}{D} = \frac{\lambda_{electron}/2}{d} \tag{5}$$

Equating the values of θ from equations (5) and (4), we get

$$\theta = \frac{\lambda_{electron}}{2d} = \frac{\lambda_{electron}}{\lambda_{photon}} \tag{6}$$

"Look!" Bill says, "$\lambda_{electron}$ cancels and we are left with "

$$\boxed{\lambda_{photon} = 2d} \tag{7}$$

"I told you," Jill interrupts, "that you had to be careful about what wavelength photons John could use. Here we see that if John's photons have a wavelength of 2d or less, his photons will carry enough of a punch, enough momentum to destroy the two slit pattern. Be sure John's photons have a wavelength longer than 2d so that they will be incapable of knocking an electron from a maxima to a minima."

"No way," responds Bill. "A wavelength of 2d is already too big. John cannot use photons with a wavelength any greater than the slit width d if he wants to see which slit the electron went through. And you want him to use photons with a wavelength greater than 2d!"

"That's the dilemma," the professor replies. "If John uses photons whose wavelength is short enough to see which slit the electron went through, he is likely to mess up the experiment and destroy the two slit pattern."

"It looks like the very act of getting information is messing up the experiment," Jill muses.

"It messes it up if we use photons," Bill responds." Let us work out a better experiment where we do a more delicate measurement to see which slit the electron went through. Do the experiment so delicately that we do not affect the motion of the electron, but accurately enough to see which slit the electron went through."

"How would you do that?" Jill asks.

"Maybe I would put a capacitor plate on one of the slits," Bill responds, "and record the capacitor voltage. If the electron went through that slit, the electric field of the electron should affect the voltage on the capacitor and leave a blip on my oscilloscope screen. If I don't see a blip, the electron went through the other slit."

"Would this measurement affect the motion of the electron?" Jill asks.

"I don't see why," Bill responds.

"Think about this," the professor interrupts. "We are now interpreting the electric and magnetic fields of a light wave as a probability wave for photons. In this view, all electric and magnetic phenomena are ultimately caused by photons. The electric and magnetic fields we worked with earlier in the course are now to be thought of as a way of describing the behavior of the underlying photons."

"That's crazy," Bill argues. "You mean, for example, that the good old $1/r^2$ Coulomb force law that holds the hydrogen atom together, is caused by photons? I don't see how."

"It's hard to visualize, but you can use a photon picture to explain every detail of the interaction between the electron and proton in the hydrogen atom. That calculation was actually done back in 1947. The modern view is that all electric and magnetic phenomena are caused by photons."

"If all electric and magnetic phenomena are caused by photons," Jill observes, "then Bill's capacitor plate and voltmeter, which uses electromagnetic phenomena, is based on photons. Since photons obey the de Broglie relationship, the photons in Bill's experiment should have the same effect as the photons from John's flashlight. If John's photons mess up the experiment, Bill's should too!"

"I have an idea," Bill says. "Aren't there such a thing as gravitational waves?"

"Yes," replies the professor. "They are very hard to make, and very hard to detect. We have not been able to make or detect them yet in the laboratory. But back in the 1970's Joe Taylor at the University of Massachusetts discovered a pair of binary neutron stars orbiting about each other. Since the stars eclipse each other, Taylor could accurately measure the orbital period."

"According to Einstein's theory of gravity, the orbiting neutron stars should radiate gravitational waves and lose energy. Joe Taylor has conclusively shown that the pair of stars are losing energy just as predicted by Einstein's theory. Taylor got the Nobel prize for this work in 1993."

"Is Einstein's theory a quantum theory?" Bill asks.

"What do you mean by that?" Jill asks.

"I mean," Bill responds, "in Einstein's theory, do gravitational waves have a particle wave nature like electromagnetic waves? Are there particles in a gravitational wave like there are photons in a light wave?"

"Not in Einstein's theory," the professor replies. "Einstein's theory is strictly a classical theory. No particles in the wave."

"Then if Einstein's theory is correct," Bill continues, "I should be able to make a gravitational wave with a very short wavelength and very little energy."

"Couldn't I then use this short wavelength, low energy, gravitational wave to see which slit the electron went through? I would make the wavelength much shorter than the slit spacing d so that there would be no doubt about which slit the electron went through. But I would use a very low energy, delicate wave so that I would not affect the motion of the electron."

"You could do that if Einstein's theory is right," the professor replies.

"But," Bill responds, "that allows me to tell which slit the electron went through without destroying the two slit pattern. What happens to the probability interpretation of the electron wave? If I know which slit the electron went through, the probability wave must have come from that slit, and we must get a one slit pattern. If John used gravitational waves instead of light waves in his flashlight, he could observe which slit the electron went through without destroying the two slit pattern."

"You have just stumbled upon one of the major outstanding problems in physics," the professor replies. "As far as we know there are four basic forces in nature. They are gravity, the electromagnetic force, the weak interaction, and the so-called gluon force that holds quarks together. I listed these in the order in which they were discovered."

"Now three of these forces, all but gravity, are known to have a particle-wave nature like light. All the particles obey the de Broglie relation $p = h/\lambda$."

"As a result, if we perform our two slit electron experiment, trying to see which slit the electron went through, and we use apparatus based on non gravitational forces, we run into the same problem we had with John's flashlight. The only chance we have for detecting which slit the electron went through without messing up the two slit pattern, is to use gravity."

"Could Einstein be wrong?" Jill asks. "Couldn't gravitational waves also have a particle nature? Couldn't the gravitational particles also obey the de Broglie relation?"

"Perhaps," the professor replies. "For years, physicists have speculated that gravity should have a particle-wave nature. They have even named the particle -- they call it a ***graviton***. One problem is that gravitons should be very, very hard to detect. The only way we know that gravitational waves actually exist is from Joe Taylor's binary neutron stars. There are various experiments designed to directly observe gravitational waves, but no waves have yet been seen in these experiments."

"In the case of electromagnetism, we saw electromagnetic radiation -- i.e., light -- long before photons were detected in Hertz's photoelectric effect experiment. After gravitational waves are detected, then we will have to do the equivalent of a photoelectric effect experiment for gravity in order to see the individual gravitons. The main problem here is that the gravitational radiation we expect to see, like that from massive objects such as neutron stars, is very low frequency radiation. Thus we would be dealing with very low energy gravitons which would be hard to detect individually."

"And there is another problem," the professor continues, "no one has yet succeeded in constructing a consistent quantum theory of gravity. There are mathematical problems that have yet to be overcome. At the present time, the only consistent theory of gravity we have is Einstein's classical theory."

"It looks like two possibilities," Jill says. "If the probability interpretation of electron waves is right, then there has to be a quantum theory of gravity, gravitons have to exist. If Einstein's classical theory is right, then there is some flaw in the probability interpretation."

"That is the way it stands now," the professor replies.

THE UNCERTAINTY PRINCIPLE

We have just seen that, for the probability interpretation of particle-waves to be a viable theory, there can be no way we can detect which slit the electron went through without destroying the two slit pattern. Also we have seen that if every particle and every force have a particle wave nature obeying the de Broglie relationship $\lambda = h/p$, then there is no way we can tell which slit the electron went through without destroying the two slit pattern. Both the particle-wave nature of matter, and the probability interpretation of particle waves, lead to a basic limitation on our ability to make experimental measurements. This basic limitation was discovered by Werner Heisenberg shortly before Schrödinger developed his wave equation for electrons. Heisenberg called this limitation the ***uncertainty principle***.

When you cannot do something, when there is really no way to do something, physicists give the failure a name and call it a basic law of physics. We began the text with the observation that you cannot detect uniform motion. Michaelson and Morley thought they could, repeatedly tried to do so, and failed. This failure is known as the principle of relativity which Einstein used as the foundation of his theories of relativity. Throughout the text we have seen the impact of this simple idea. When combined with Maxwell's theory of light, it implied that light traveled at the same speed relative to all observers. That implied moving clocks ran slow, moving lengths contracted, and the mass of a moving object increased with velocity. This led to the relationship $E = mc^2$ between mass and energy, and to the connections between electric and magnetic fields. The simple idea that you cannot measure uniform motion has an enormous impact on our understanding of the way matter behaves.

Now, with the particle-wave nature of matter, we are encountering an equally universal restriction on what we can measure, and that restriction has an equally important impact on our understanding of the behavior of matter. Our discussion of the uncertainty principle comes at the end of the text rather than at the beginning only because it has taken a while to develop the concepts we need to explain this restriction. With the principle of relativity we could rely on the student's experience with uniform motion, clocks and meter sticks. For the uncertainty principle, we need some understanding of the behavior of particles and waves, and as we shall see, Fourier analysis plays an important role.

There are two forms of the uncertainty principle, one related to measurements of position and momentum, and the other related to measurements of time and energy. They are not separate laws, one can be derived from the other. The choice of which to use is a matter of convenience. Our discussion of the two slit experiment and the de Broglie relationship naturally leads to the position-momentum form of the law, while Fourier analysis naturally introduces the time-energy form.

POSITION-MOMENTUM FORM OF THE UNCERTAINTY PRINCIPLE

In our two slit thought experiment, in the attempt to see which slit the electron went through, we used a beam of photons whose momenta was related to their wavelength by $p = h/\lambda$. The wave nature of the photon is important because we cannot see details smaller than a wavelength λ when we scatter waves from an object. When we use waves of wavelength λ, the uncertainty in our measurement is at least as large as λ. Let us call the uncertainty in the position measurement Δx.

However when we use photons to locate the electron, we are slugging the electron with particles, photons of momentum $p_{photon} = h/\lambda$. Since we do not know where the photons are within a distance λ, we do not know exactly how the electron was hit and how much momentum it absorbed from the photon. The electron could have absorbed the full photon momentum p_{photon} or none of it. If we observe the electron, we make the electron's momentum uncertain by an amount at least as large as p_{photon}. Calling the uncertainty in the electron's momentum $\Delta p_{electron}$ we have

$$\Delta p_{electron} = p_{photon} = \frac{h}{\lambda} = \frac{h}{\Delta x} \tag{8}$$

multiplying through by Δx gives

$$\Delta p \Delta x = h \tag{9}$$

In Equation 9, Δp and Δx represent the smallest possible uncertainties we can have when measuring the position of the electron using photons. To allow for the fact that we could get much greater uncertainties using poor equipment or sloppy techniques, we will write the equation in the form

$$\boxed{\Delta p \Delta x \geq h} \qquad \textit{position–momentum form of the uncertainty principle} \tag{10}$$

indicating that the product of the uncertainties is at least as large as Planck's constant h.

If all forces have a particle nature, and all particles obey the de Broglie relationship, then the fact that we derived Equation 10 using photons makes no difference. We have to get the same result using any particle, in any possible kind of experiment. Thus Equation 10 represents a fundamental limitation on the measurement process itself!

Equation 10 is not like any formula we have previously dealt with in the text. It gives you an estimate, not an exact value. Often you will see the formula written $\Delta p \Delta x \geq \hbar$ with $\hbar = h/2\pi$, rather than h, appearing on the right side. Whether you use h or $\hbar$ depends upon how you wish to define the uncertainties Δp and Δx. But it is not necessary to be too precise. The important point is that the product $\Delta p \Delta x$ must be at least of the order of magnitude h. It cannot be h/100 or something smaller.

The gist of the uncertainty principle is that the more accurately you measure the position of the particle, the more you mess up the particle's momentum. Or, the more accurately you measure the momentum of a particle, the less you know about the particle's position.

Equation (10) is not quite right, because it turns out that an accurate measurement of the x position of a particle does not necessarily mess up the particle's y component of momentum, only its x component. A more accurate statement of the uncertainty principle is

$$\Delta p_x \Delta x \geq h \tag{11a}$$

$$\Delta p_y \Delta y \geq h \tag{11b}$$

where Δp_x is the uncertainty in the particle's x component of momentum due to a measurement of its x position, and Δp_y is the uncertainty in the y component of momentum resulting from a y position measurement. The quantities Δx and Δy are the uncertainty in the x and y measurements respectively.

Single Slit Experiment

In our two slit thought experiment, we measured the position of the electron by hitting it with a photon. Another way to measure the position of a particle is to send it through a slit. For example, suppose a beam of particles impinges on a slit of width (w) as illustrated in Figure (11). We know that any particle that makes it to the far side of the slit had, at one time, been within the slit. At that time we knew its y position to within an uncertainty Δy equal to the width (w) of the slit.

$$\Delta y = w \tag{12}$$

This is an example of a position measurement with a precisely known uncertainty Δy .

According to the uncertainty principle, the particle's y component of momentum is uncertain by an amount Δp_y given by Equation 11b as

$$\Delta p_y \ge \frac{h}{\Delta y} = \frac{h}{w} \tag{13}$$

Equation 13 tells us that the smaller Δy , i.e., the narrower the slit, the bigger the uncertainty Δp_y we create in the particle's y momentum. This is what happens if the particle's motion is governed by its wave nature.

In Figure (12a) we have a ripple tank photograph of a wave passing through a moderately narrow slit. The wave on the far side of the slit is seen to spread out a bit. We can calculate the amount of spread by noting that the beam is mostly contained within the central maximum of the single slit diffraction pattern.

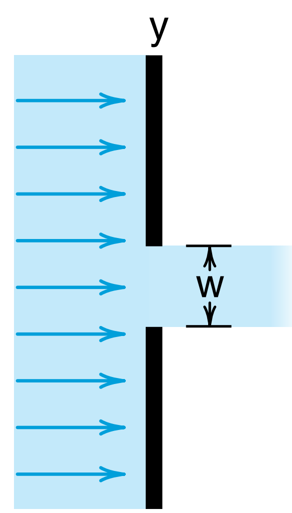


Figure 11
When a particle goes through the slit, its y position is known to within an uncertainty $\Delta y = w$.

Now suppose that this wave represented a beam of photons or electrons. On the right side of the slit, all the particles have a definite x component of momentum $p_x = h/\lambda$ and no y momentum. The uncertainty in the y momentum is zero.

Once the particle's have gone through the slit, the beam spreads out giving the particles a y momentum.

Since you do not know whether any given particle in the beam will go straight ahead, up or down, the spread of the beam introduces an uncertainty Δp_y in the particle's y momentum. This spread is illustrated in Figure (12b).

In Figure (13a), we see a wave passing through a narrower slit than the one in Figure (12a). With the narrower slit, we have made a more precise measurement of the particle's y position. We have reduced the uncertainty $\Delta y = w$. According to the uncertainty principle $\Delta p_y \geq h/\Delta y$, a decrease in Δy should increase the uncertainty Δp_y in the particle's y momentum. But an increase in Δp_y means that the beam should spread out more, which is what it does in Figure (13). In going from Figure (12) to Figure (13), we have cut the slit width about in half and about doubled the spread. I.e., cutting Δy in half doubles Δp_y as expected.

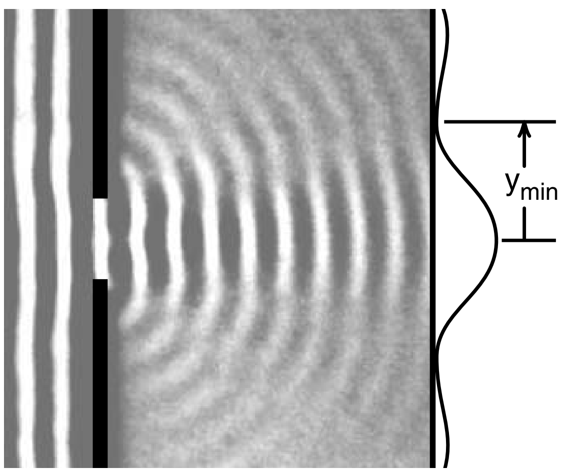


Figure 12a
Wave picture of the single slit experiment, wide slit.

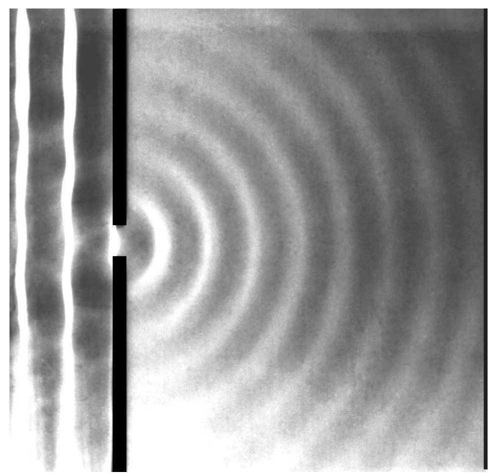


Figure 13a
Narrow the slit and the wave spreads out.

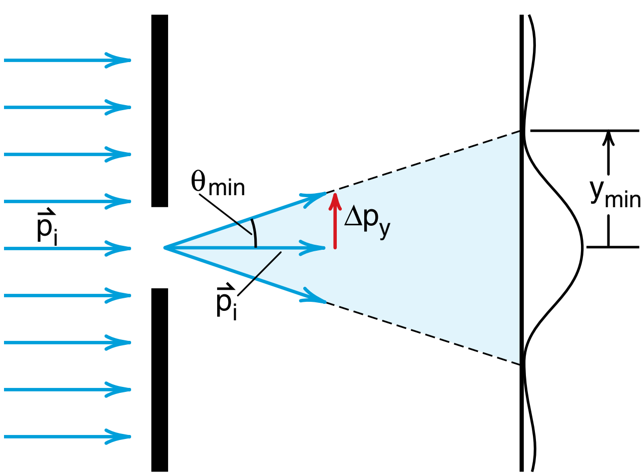


Figure 12b
Particle picture of the single slit experiment.

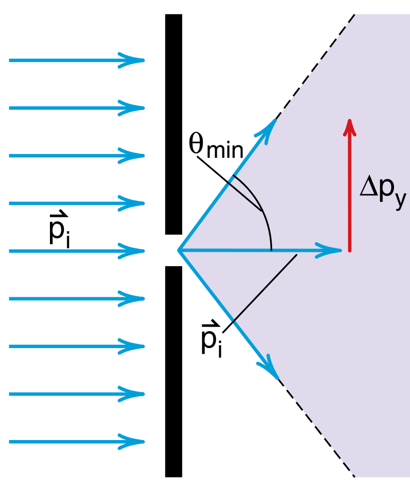


Figure 13b
With a narrower slit, Δp_y increases.

Example 1

We can use our analysis of the single slit pattern in Chapter 33 to show that Δp_y and Δy are related by the uncertainty principle. We saw that when a wave goes through a single slit of width w, the distance y_{min} to the first minimum is given by

$$y_{min} = \frac{\lambda D}{w} \qquad (33\text{-}14)$$

where λ is the wavelength and D the distance to the screen as shown in Figure (14a). The angle to the first minimum is given by

$$\tan(\theta_{min}) = \frac{y_{min}}{D} = \frac{(\lambda D/w)}{D} = \frac{\lambda}{w} \qquad (14)$$

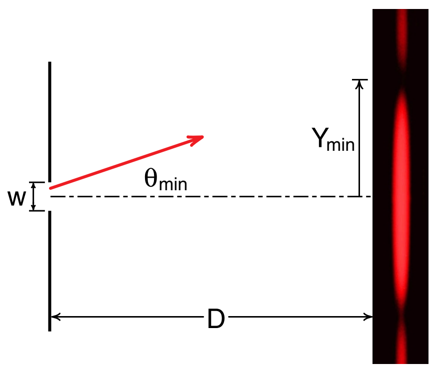


Figure 14a

After the beam emerges from the slit, the momenta of the particles spreads out through the same angle θ_{min} as indicated in Figure (14b). From that figure we have

$$\tan(\theta_{min}) = \frac{\Delta p_y}{p_x} \qquad (15)$$

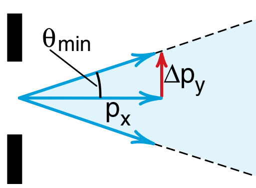


Figure 14b

where Δp_y represents the possible spread in the particle's y momenta. Equating values of $\tan(\theta_{min})$ from Equation 14 and Equation 15 gives

$$\frac{\lambda}{w} = \frac{\Delta p_y}{p_x} \qquad (16)$$

The particles entered the slit as plane waves with only an x component of momentum given by de Broglie's formula

$$p_x = \frac{h}{\lambda} \qquad (17)$$

Using this formula for p_x in Equation 16 gives

$$\frac{\lambda}{w} = \frac{\Delta p_y}{(h/\lambda)} = \frac{\lambda \Delta p_y}{h} \qquad (18)$$

The λ's cancel, and we are left with

$$(\Delta p_y)w = h \qquad (19)$$

But the slit width w is Δy, the uncertainty in the y measurement, thus

$$\Delta p_y \Delta y = h \qquad (20)$$

There is an equal sign in Equation 20 because this particular measurement of the y position of the particle causes the least possible uncertainty in the particle's y component of momentum. (Note that the x component of the particle's momentum is more or less unaffected by the slit. The wave has the same wavelength λ before and after going through the slit. It is the y component of momentum that changed from zero on the left side to $\pm\Delta p_y$ on the right.)

Exercise 3

A microwave beam, consisting of 1.24×10^{-4}eV photons impinges on a slit of width (w) as shown in Figure 14a.

(a) What is the momentum p_x of the photons in the laser beam before they get to the slit?

(b) When the photon passes through the slit, their y position is known to an uncertainty $\Delta y = w$, the slit width. Before the photons get to the slit, their y momentum has the definite value $p_y = 0$. Passing through the slit makes the photon's y momentum uncertain by an amount Δp_y. Using the uncertainty principle, calculate what the slit width (w) must be so that Δp_y is equal to the photon's original momentum p_x. How does w compare with the wavelength λ of the laser beam?

(c) If Δp_y becomes as large as the original momentum p_x, what can you say about the wave pattern on the right side of Figure (11)? Is this consistent with what you know about waves of wavelength λ passing through a slit of this width? Explain.

TIME-ENERGY FORM OF THE UNCERTAINTY PRINCIPLE

The second form of the uncertainty principle, which perhaps has an even greater impact on our understanding of the behavior of matter, involves the measurement of the energy of a particle, and the time available to make the measurement. ***The shorter the time available, the less accurate the energy measurement is***. If ΔE is the uncertainty in the results of our energy measurement, and Δt the time we had to make the measurement, then ΔE and Δt are related by

$$\boxed{\Delta E \Delta t \ \geq\ h} \tag{21}$$

One can derive this form of the uncertainty principle from $\Delta p \Delta x \ \geq\ h$, but we can gain a better insight into the relationship by starting with an explicit example.

A device that has become increasingly important in research, particularly in the study of fast reactions in molecules and atoms, is the pulsed laser. The lasers we have used in various experiments are all continuous beam lasers. The beam is at least as long as the distance from the laser to the wall. If we had a laser that we could turn on and off in one nanosecond, the pulse would be 1 foot or 30 cm long and contain

$$\frac{30\text{cm}}{6\times10^{-5}\dfrac{\text{cm}}{\text{wavelength}}} = 5\times10^{5}\text{wavelengths}$$

Even a picosecond laser pulse which is 1000 times shorter, contains 500 wavelengths. Some of the recent pulsed lasers can produce a pulse 500 times shorter than that, only 2 femtoseconds (2×10^{-15} seconds) long. These lasers emit a pulse that is only one wavelength long.

For our example of the time-energy form of the uncertainty principle, we wish to consider the nature of the photons in a 2 femtosecond long laser pulse. If we want to measure the energy of the photons in such a pulse, we only have 2 femtoseconds to make the measurement because that is how long the pulse takes to go by us. In the notation of the uncertainty principle

$$\begin{aligned}\Delta t &= 2\times10^{-15}\text{sec} \\ &= 2\text{ femtoseconds}\end{aligned} \qquad \begin{array}{l}\textit{time available to}\\ \textit{measure the energy}\\ \textit{of the photons}\\ \textit{in our laser pulse}\end{array} \tag{22}$$

Let us suppose that the laser produces red photons whose wavelength is $6.2\times10^{-5}\text{cm}$, about the wavelength of the lasers we have been using. According to our usual formula for calculating the energy of the photons in such a laser beam we have

$$E_{\text{photon}} = \frac{12.4\times10^{-5}\text{eV cm}}{6.2\times10^{-5}\text{cm}} = 2\text{ eV} \tag{23}$$

Now let us use the uncertainty principle in the form

$$\Delta E \ \geq\ \frac{h}{\Delta t} \tag{24}$$

to calculate the uncertainty in any measurement we would make the energy of the photons in the 2 femtosecond laser beam. We have

$$\begin{aligned}\Delta E \ \geq\ \frac{h}{\Delta t} \ &\geq\ \frac{6.63\times10^{-27}\text{erg sec}}{2\times10^{-15}\text{sec}} \\ &\geq\ 3.31\times10^{-12}\text{ergs}\end{aligned} \tag{25}$$

Converting ΔE from ergs to electron volts, we get

$$\Delta E \ \geq\ \frac{3.31\times10^{-12}\text{ergs}}{1.6\times10^{-12}\text{ergs/eV}} \approx 2\text{ eV} \tag{26}$$

The uncertainty in any energy measurement we make of these photons is as great as the energy itself! If we try to measure the energy of these photons, we expect the answers to range from $E - \Delta E = 0\text{ eV}$ up to $E + \Delta E = 4\text{ eV}$. Why does this happen? Why is the energy of the photons in this beam so uncertain? ***Fourier analysis provides the answer***.

We can see why the energy of the photons in the 2 femtosecond pulse is so uncertain by comparing the Fourier transform of a long laser pulse with that of a pulse consisting of only one wavelength.

Figure (15) shows the Fourier transform of an infinitely long sine wave. You will recall that, in the design of the MacScope program, it is assumed that we are analyzing a repeated waveform. If you continuously repeat the waveform seen in the upper half of the diagram, you get an infinitely long cyclic wave which is a pure sine wave. (Sine waves are by definition infinitely long waves.) In effect we have in Figure (15) selected 16 cycles of the pure sine wave, and the Fourier analysis box shows that we have a pure 16th harmonic. This sine wave has a definite frequency f, and if this represented a laser beam, the photons in the beam would have a precise energy given by the formula E = hf. There is no uncertainty in the energy of this infinitely long sine wave. (It would take an infinite time Δt to make sure that the wave was infinitely long, with the result $\Delta E = h/\Delta t = h/\infty = 0$.)

In Figure (16) we are looking at a waveform consisting of a single pulse. This would accurately represent the output of a red laser that continuously emitted single wavelength pulses spaced 16 wavelengths apart. (Remember that our program assumes that the wave shape is repeated.) From the Fourier analysis box we see that there is a dramatic difference between the composition of a pure sine wave and of a single pulse. To construct a single pulse out of sine waves, we have to add up a slew of harmonics. The single pulse is more like a drum beat while the continuous wave is more like a flute. (In the appendix we show how the sine wave harmonics add up to produce a pulse.)

In Figure (16) we see that the dominant harmonic is still around the sixteenth, as it was for the continuous wave, but there is a spread of harmonics from near zero up to almost the 32nd. For a laser pulse to have this shape, it must consist of frequencies ranging from near zero up to twice the natural frequency. Each of these frequencies contains photons whose energy is given by Einstein's formula E = hf where f is the frequency of the harmonic.

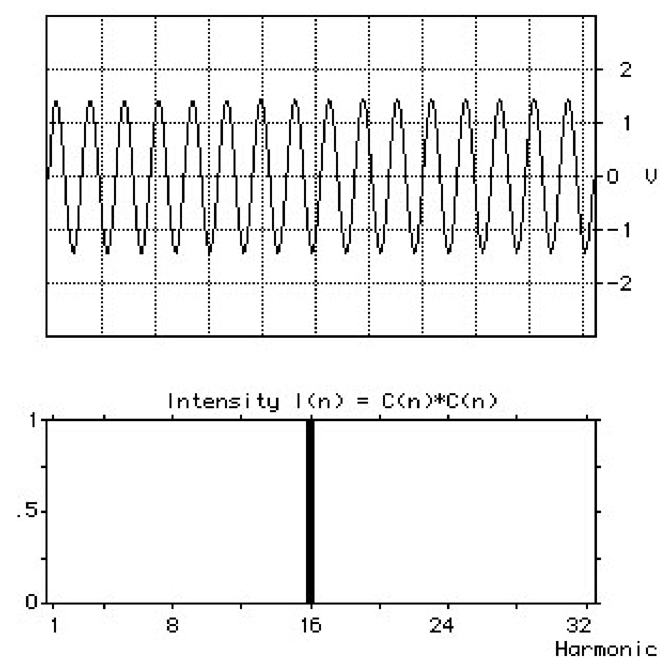


Figure 15
A pure sine wave has a single frequency.

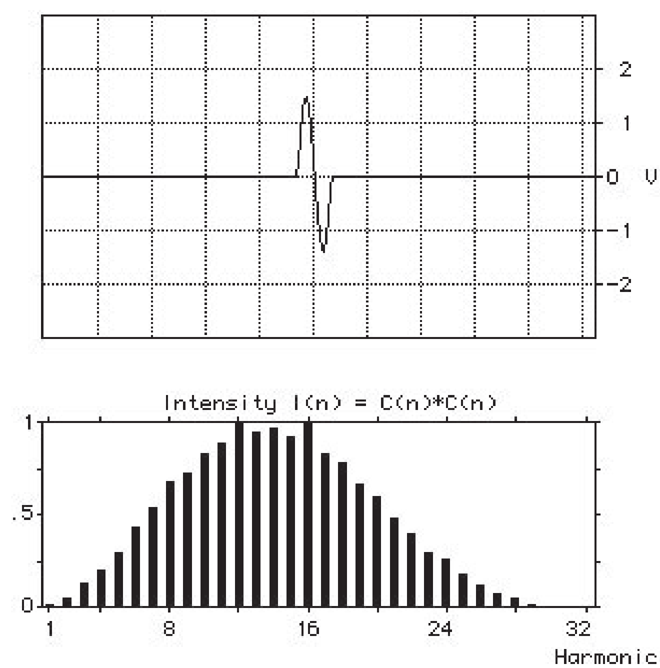


Figure 16
One cycle of a wave is made up of a spread of harmonics

In Figure (17) we have reproduced the Fourier analysis box of Figure (16), but relabeled the horizontal axis in electron volts. We have assumed that the 16th harmonic represented 2 eV photons which would be the case if the wave were an infinitely long red laser pulse. Now the diagram represents the density of photons of different energies in the laser beam. While 2 eV is the most likely energy, there is a spread of energies ranging from nearly 0 eV up to nearly 4 eV. If we measure the energy of a photon in the beam, our answer is 2 eV with an uncertainty of 2 eV, just as predicted by the uncertainty principle $\Delta E \Delta t \geq h$.

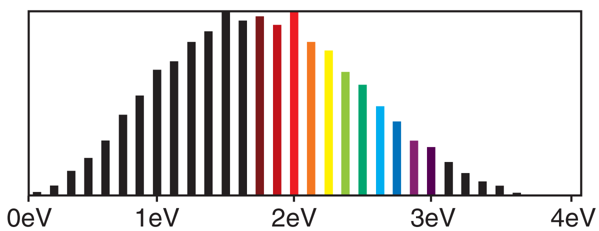


Figure 17
Photon energies in single wavelength pulse of a red laser beam.

In Figure (18) we analyze a pulse two wavelengths long. Now we see that the spread of frequencies required to reconstruct this waveform is only half as wide, ranging from the 8th to 24th harmonic, or from 1 eV to 3 eV. We have twice as long to study a 2 wavelength pulse, and the uncertainty in energy ΔE is only about 1 eV, or half as big.

Going to a 4 wavelength pulse in Figure (19) we see that by doubling the time available we again cut in half the uncertainty ΔE in energy. Now the energy varies from about 1.5 eV to 2.5 eV for a $\Delta E = .5$ eV . This is just what you expect from $\Delta E \Delta t \geq h$. You should now begin to see that the uncertainty principle is a simple rule evolving from the wave nature of particles. (By the way, it would be more accurate to write $\Delta E \Delta t = h$ for this discussion, because we are describing the very least uncertainty in energy.)

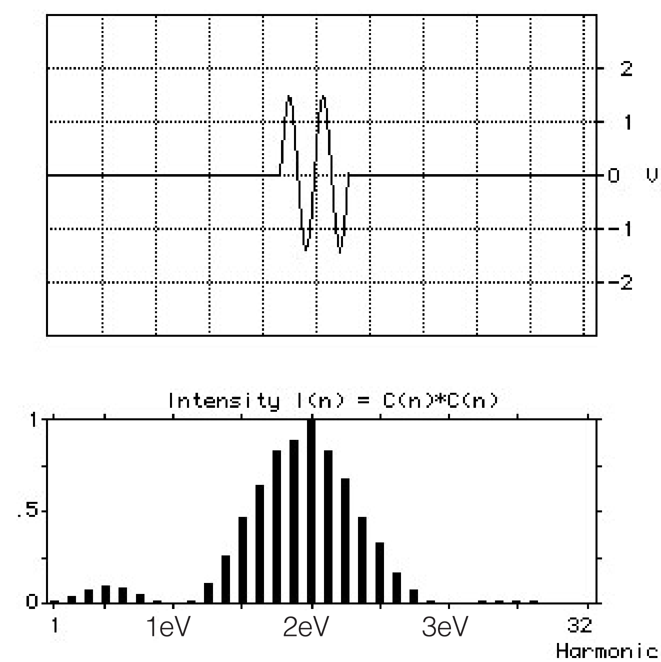


Figure 18
A two cycle wave has half the spread of harmonics.

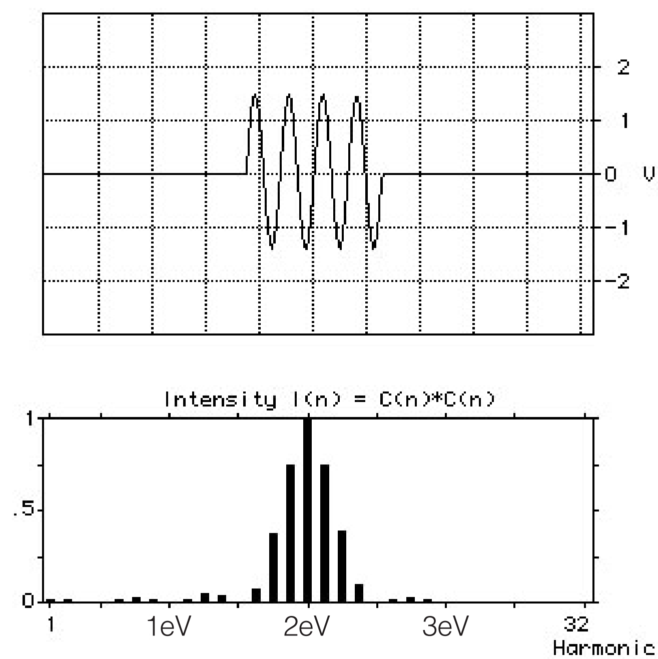


Figure 19
A four cycle wave has a fourth the spread of harmonics.

Probability Interpretation

We have interpreted Figure (17) as representing the spread in energies of the photons in a 2 femtosecond red laser pulse. What if the pulse consisted of only a single photon? Then how do we interpret this spread in energies? The answer is that we use a probability interpretation. The photon in the pulse has different probabilities of having different energies.

In our discussion of light waves, we saw that the energy density in a light wave was proportional to the square of the amplitude of the wave. This is reasonable because while the amplitude of a wave can be positive or negative, the square of the amplitude, which we call the ***intensity*** is always positive. Probabilities, like energy densities, also have to be positive, thus we should associate the probability of a photon as having a given frequency with the intensity or square of the amplitude of the wave of that frequency.

In Figure (20) we show the intensities (square of the amplitudes) of the harmonics that make up the single wavelength pulse. (This is plotted automatically by MacScope when we click on the button labeled Φ.) We see that squaring the amplitudes narrows the spread.

Figure (20) has the following interpretation when applied to pulses containing a single photon. If we measure the energy of the photon, we are most likely to get an answer close to 2 eV but there is a reasonable probability of getting an answer lower than 1 eV or even higher than 3 eV. The heights of the bars tell us the relative probability of measuring that energy for the photon.

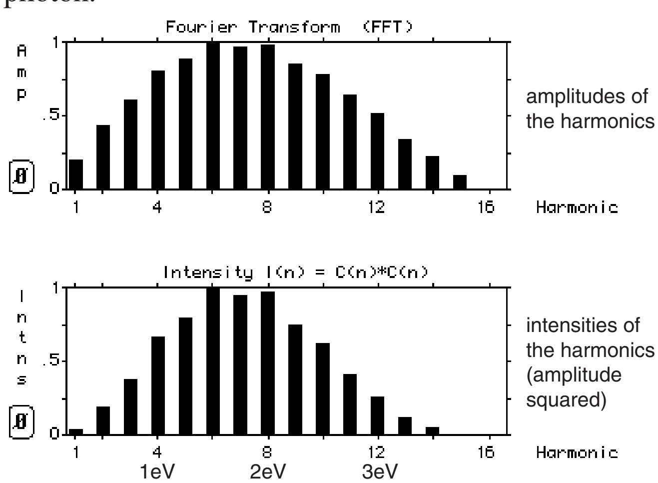


Figure 20
Intensities of the harmonics are proportional to the square of the amplitudes.

Measuring Short Times

We have said that the new pulsed lasers produce pulses as short as 2 femtoseconds. How do we know that? Suppose we gave you the job of measuring the length of the laser pulse, and the best oscilloscope you had could measure times no shorter than a nanosecond. This is a million times too slow to see a femtosecond pulse. What do you do?

If you cannot measure the time directly, you can be sneaky and use the uncertainty principle. Send the laser pulse through a diffraction grating, and record the spread in wavelengths, i.e., the spread in energies of the photons in the pulse. If the line is very sharp, if they are all red photons of a single wavelength and energy, then you know that there is no measurable uncertainty ΔE in the photon energies, and the pulse must last a time Δt that is considerably longer than 2 femtoseconds. If, on the other hand, the line is spread out from the near infra red to violet, if the spread in energies is from 1 eV to 3 eV, and the spread is not caused by some other phenomena (like the Doppler effect), then from the uncertainty principle you know that the pulse is only about a femtosecond long. (You know, for example, it cannot be as long as 10 femtoseconds, or as short as a tenth of a femtosecond.)

Thus, with the uncertainty principle, you can use a diffraction grating rather than a clock or oscilloscope to measure very short times. Instead of being an annoying restriction on our ability to make experimental measurements, the uncertainty principle can be turned into an important scientific tool for measuring short times and, as we shall see, short distances.

Exercise 4

An electron is in an excited state of the hydrogen atom, either the second energy level at -3.40 eV or the third energy level at -1.51 eV. You wish to do an experiment to decide which of these two states the electron is in. What is the least amount of time you **must** take to make this measurement?

Short Lived Elementary Particles

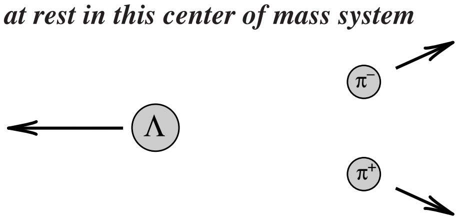
We usually think of the rest energy of a particle as having a definite value. For example the rest energy of a proton is $938.2723 \times 10^6 \text{eV}$. The proton itself is a composite particle made of 3 quarks, and the number 938.2723 MeV represents the total energy of the quarks in the allowed wave pattern that represents a proton. This rest energy has a very definite value because the proton is a stable particle with plenty of time to settle into a precise wave pattern.

A rather different particle is the so called "$\Lambda(1520)$", which is another combination of 3 quarks, but very short lived. The name comes partly from the fact that the particle's rest mass energy is about 1520 million electron volts (MeV). As indicated in Figure (21), a $\Lambda(1520)$ can be created as a result of the collision between a K^- meson and a proton.

We are viewing the collision in a special coordinate system, where the total momentum of the incoming particles is zero. In this coordinate system, the resulting $\Lambda(1520)$ will be at rest. By conservation of energy, the total energy of the incoming particles should equal the rest mass energy of the $\Lambda(1520)$. Thus if we collide K^- particles with protons, we expect to create a $\Lambda(1520)$ particle only if the incoming particles have the right total energy.

a) A K^- meson and a proton are about to collide. We are looking at the collision in a coordinate system where the total momentum is zero (the so called "center of mass" system).

b) ***In the collision a $\Lambda(1520)$ particle is created. It is at rest in this center of mass system***



c) ***The $\Lambda(1520)$ then quickly decays into a lower energy Λ particle and two π mesons.***

Figure 21
A $\Lambda(1520)$ particle can be created if the total energy (in the center of mass system) of the incoming particles equals the rest mass energy of the $\Lambda(1520)$.

Figure (22) shows the results of some collision experiments, where a K^- meson and a proton collided to produce a Λ and two π mesons. The probability of such a result peaked when the energy of the incoming particles was 1,520 MeV. This peak occurred because the incoming K^- meson and proton created a $\Lambda(1520)$ particle, which then decayed into the Λ and two π mesons, as shown in Figure (21). The $\Lambda(1520)$ was not observed directly, because its lifetime is too short.

Figure (22) shows that the energy of the incoming particles does not have to be exactly 1520 MeV in order to create a $\Lambda(1520)$. The peak is in the range from about 1510 to 1530 Mev, which implies that the rest mass energy of the $\Lambda(1520)$ is 1520 MeV plus or minus about 10 MeV. From one experiment to another, the rest mass energy can vary by about 20 MeV. (The experimentalists quoted a variation of 16 MeV.)

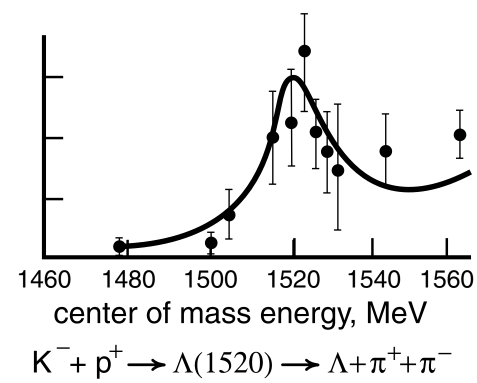


Figure 22
The probability that a K^- meson and a proton collide to produce a Λ particle, and two π mesons peak at an energy of 1520 MeV. The peak results from the fact that a $\Lambda(1520)$ particle was created and quickly decayed into the Λ and two π mesons. The probability peaks at 1520 MeV, but can be seen to spread out over a range of about 16MeV. The small circles are experimental values, the vertical lines represent the possible error in the value. (Data from M.B. Watson et al., **Phys. Rev.** ***131(1963).)***

Why isn't the peak sharp? Why does the rest mass energy of the $\Lambda(1520)$ particle vary by as much as 16 to 20 MeV from one experiment to another? The answer lies in the fact that ***the lifetime of the $\Lambda(1520)$ is so short, that the particle does not have enough time to establish a definite rest mass energy***. The 16 MeV variation is the uncertainty ΔE in the particles rest mass energy that results from the fact that the particle's lifetime is limited.

The uncertainty principle relates the uncertainty in energy ΔE to the time Δt available to establish that energy. To establish the rest mass energy, time Δt available is the particle's ***lifetime***. Thus we can use the uncertainty principle to estimate the lifetime of the $\Lambda(1520)$ particle. With $\Delta E \times \Delta t \approx h$ we get

$$\Delta t \approx \frac{h}{\Delta E} = \frac{6.63 \times 10^{-27}\,\text{erg sec}}{16\ \text{MeV} \times \left(1.6 \times 10^{-6}\frac{\text{erg}}{\text{MeV}}\right)}$$

$$\Delta t \approx 2.6 \times 10^{-22}\ \text{seconds} \tag{27}$$

The lifetime of the $\Lambda(1520)$ particle is of the order of 10^{-22} seconds! This is only about 10 times longer than it takes light to cross a proton! Only by using the uncertainty principle could we possibly measure such short times.

THE UNCERTAINTY PRINCIPLE AND ENERGY CONSERVATION

The fact that for short times the energy of a particle is uncertain, raises an interesting question about basic physical laws like the law of conservation of energy. If a particle's energy is uncertain, how do we know that energy is conserved in some process involving that particle? The answer is -- ***we don't***.

One way to explain the situation is to say that nature will cheat if it can get away with it. Energy does not have to be conserved if we cannot do an experiment to demonstrate a lack of conservation of energy.

Consider the process shown in Figure (23). It shows a red, 2 eV photon traveling along in space. Suddenly the photon creates an positron-electron pair. The rest mass energy of both the positron and the electron are .51 MeV. Thus we have a 2 eV photon creating a pair of particles whose total energy is 1.02×10^6eV, a huge violation of the law of conservation of energy. A short time later the electron and positron come back together, annihilate, leaving behind a 2 eV photon. This is an equally huge violation of the conservation of energy.

But have we really violated the conservation of energy? During its lifetime, the positron-electron pair is a composite object whose total energy is uncertain. If the pair lived a long time, its total energy would be close to the expected energy of 1.02×10^6eV. But suppose the pair were in existence only for a very short time Δt, a time so short that the uncertainty in the energy could be as large as 1.02×10^6eV. Then there is some probability that the energy of the pair might be only 2 eV and the process shown in Figure (2) could happen.

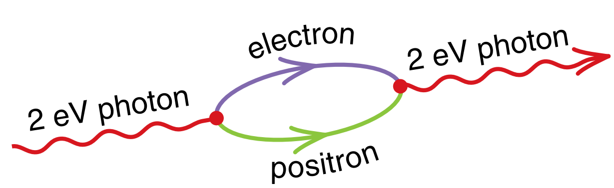


Figure 23
Consider a process where a 2 eV photon suddenly creates a positron-electron pair. A short time later the pair annihilates, leaving a 2 eV photon. In the long range, energy is conserved.

The length of time Δt that the pair could exist and have an energy uncertain by 1.02 MeV is

$$\Delta t = \frac{h}{\Delta E} = \frac{6.63 \times 10^{-27}\,\text{erg sec}}{1.02\,\text{MeV} \times 1.6\times 10^{-6}\frac{\text{erg}}{\text{MeV}}}$$

$$\Delta t = 4 \times 10^{-21}\text{sec} \tag{32}$$

Another way to view the situation is as follows. Suppose the pair in Figure (20) lasted only 4×10^{-21} seconds or less. Even if the pair had an energy of $1.02 \times 10^{6}\text{eV}$, the lifetime is so short that any measurement of the energy of the pair would be uncertain by at least $1.02 \times 10^{6}\text{eV}$, and the experiment could not detect the violation of the law of conservation of energy. In this point of view, if we cannot perform an experiment to detect a violation of the conservation law, then the process should have some probability of occurring.

Does a process like that shown in Figure (23) actually occur? If so, is there any way that we can know that it does? The answer is yes, to both questions. It is possible to make extremely accurate studies of the energy levels of the electron in hydrogen, and to make equally accurate predictions of the energy using the theory of ***quantum electrodynamics***. We can view the binding of the electron in hydrogen as resulting from the continual exchange of photons between the electron and proton. During this continual exchange, there is some probability that the photon creates a positron electron pair that quickly annihilates as shown in Figure (23). In order to predict the correct values of the hydrogen energy levels, the process shown in Figure (23) has to be included. Thus we have direct experimental evidence that for a short time the particle antiparticle pair existed.

QUANTUM FLUCTUATIONS AND EMPTY SPACE

We began the text with a discussion of the principle of relativity—that you could not detect your own motion relative to empty space. The concept of empty space seemed rather obvious—space with nothing in it. But the idea of empty space is not so obvious after all.

With the discovery of the cosmic background radiation, we find that all the space in this universe is filled with a sea of photons left over from the big bang. We can accurately measure our motion relative to this sea of photons. The earth is moving relative to this sea at a velocity of 600 kilometers per second toward the Vergo cluster of galaxies. While this measurement does not violate the principle of relativity, it is in some sense a measurement of our motion relative to the universe as a whole.

Empty space itself may not be empty. Consider a process like that shown in Figure (24) where a photon, an electron, and a positron are all created at some point in space. A short while later the three particles come back together with the positron and electron annihilating and the photon being absorbed.

One's first reaction might be that such a process is ridiculous. How could these three particles just appear and then disappear? To do this we would have to violate both the laws of conservation of energy and momentum.

But, of course, the uncertainty principle allows us to do that. We can, in fact, use the uncertainty principle to estimate how long such an object could last. The arguments would be similar to the ones we used in the analysis of the process shown in Figure (23).

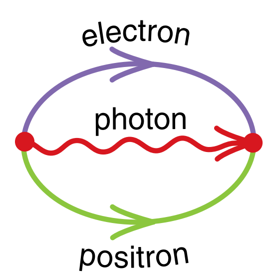


Figure 24
Quantum fluctuation. The uncertainty principle allows such an object to suddenly appear, and then disappear.

In the theory of quantum electrodynamics, a completely isolated process like that shown in Figure (24) does not affect the energy levels of the hydrogen atom and should be undetectable in electrical measurements. But such a process might affect gravity. A gravitational wave or a graviton might interact with the energy of such an object. Some calculations have suggested that such interactions could show up in Einstein's classical theory of gravity as a correction to the famous ***cosmological constant*** we discussed in Chapter 21.

An object like that shown in Figure (23) is an example of what one calls a ***quantum fluctuation***. Here we have something that appears and disappears in so-called empty space. If such objects can keep appearing and disappearing, then we have to revise our understanding of what we mean by empty.

The uncertainty principle allows us to tell the difference between a quantum fluctuation and a real particle. A quantum fluctuation like that in Figure (24) violates conservation of energy, and therefore cannot last very long. A real particle can last a long time because energy conservation is not violated.

However, there is not necessarily that much difference between a real object and a quantum fluctuation. To see why, let us take a closer look at the π meson. The π^+ is a particle with a rest mass energy of 140 MeV, that consists of a quark-antiquark pair. The quark in that pair is the so-called ***up*** quark that has a rest mass of roughly 400 MeV. The other is the ***antidown*** quark that has a rest mass of about 700 MeV. (Since we can't get at isolated quarks, the quark rest masses are estimates, but should not be too far off). Thus the two quarks making up the π meson have a total rest mass of about 1100 MeV. How could they combine to produce a particle whose rest mass is only 140 MeV?

The answer lies in the potential energy of the gluon force that holds the quarks together. As we have seen many times, the potential energy of an attractive force is negative. In this case the potential energy of the gluon force is almost as big in magnitude as the rest mass of the quarks, reducing the total energy from 1100 MeV to 140 MeV.

Suppose we had an object whose negative potential energy was as large as the positive rest mass energy. Imagine, for example, that the object consisted of a collection of point sized elementary particles so close together that their negative gravitational potential energy was the same magnitude as the positive rest mass and kinetic energy. Suppose such a collection of particles were created in a quantum fluctuation. How long could the fluctuation last?

Since such an object has no total energy, the violation ΔE of energy conservation is zero, and therefore the lifetime $\Delta t = h/\Delta E$ could be forever.

Suppose the laws of physics required that such a fluctuation rapidly expand, greatly increasing both the positive rest mass and kinetic energy, while maintaining the corresponding amount of negative gravitational potential energy. As long as ΔE remained zero, the expanding fluctuation could keep on going. Perhaps such a fluctuation occurred 14 billion years ago and we live in it now.

APPENDIX

HOW A PULSE IS FORMED FROM SINE WAVES

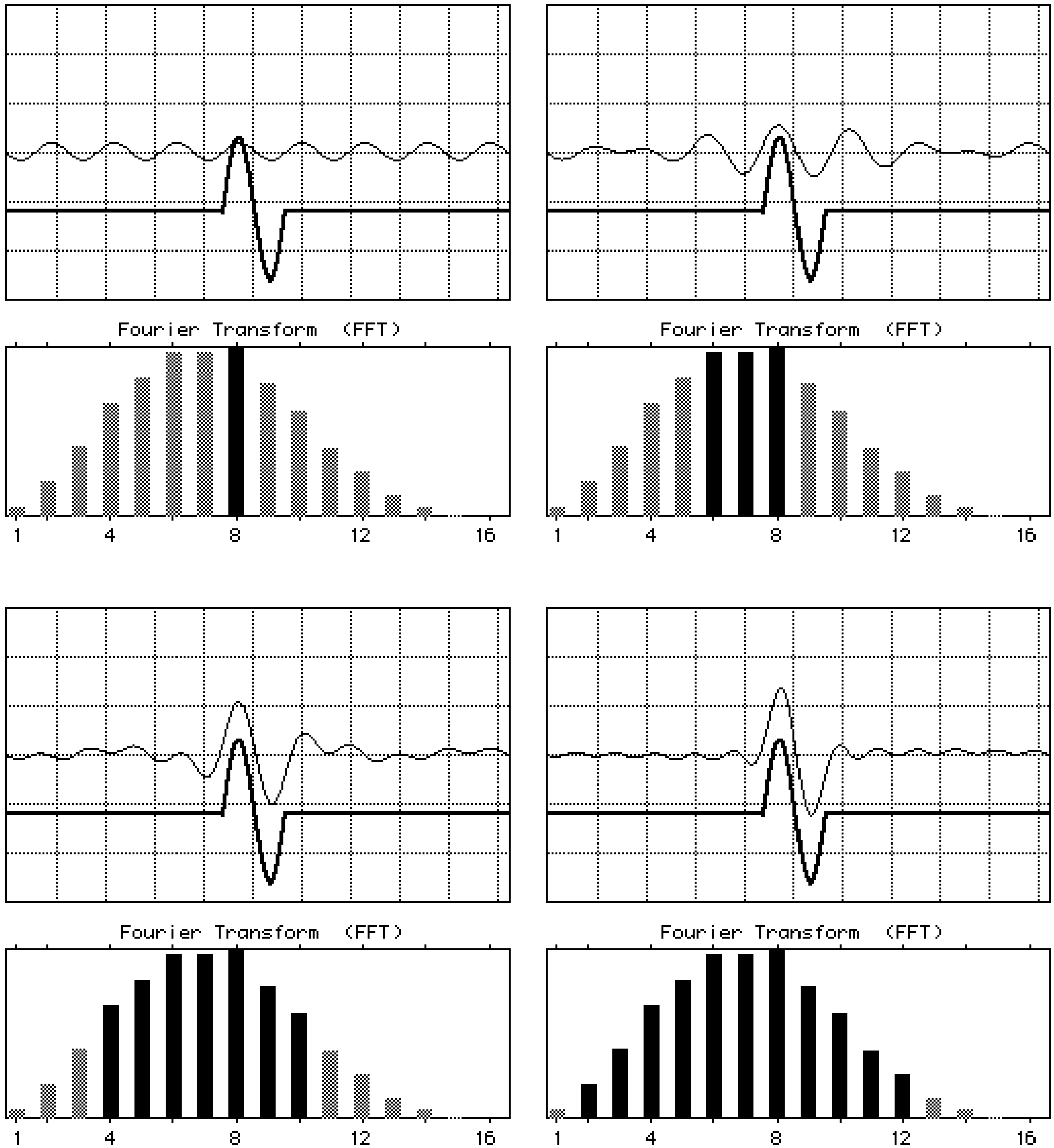


Figure A1
By selecting more and more harmonics, you can see how the sine waves add up to produce a pulse.

Index

Contents