

Foundations, 1

Quantum mechanics is money

Text message and take a picture with your smart phone; watch a movie on your Blu-ray player; get the bar code on your bag of chips scanned; obtain an MRI image of your aching shoulder; take a ride on a maglev train. None of these—and countless other—things would be possible without quantum mechanics! Leon Lederman, Nobel Prize winning physicist, is widely quoted as saying that 1/3 of the world's economy is due to quantum mechanics ($QM = \$$). Lederman's estimate is actually probably too low, but what surely is the case is that computers, lasers, and superconducting magnets (to cite a just few examples), and all of the very familiar products that rely on them, could not exist in their present forms without knowledge of the quantum mechanics of electrons and photons. A poignant demonstration of this is found in the image to the right. The large crate being “uploaded” to the Pan Am plane is 5 MB of computer memory. The flash drive shown would be only a dot on the side of the crate yet contains 1 Tb of memory—equivalent to about 200,000 crates (and, in today's dollars, 10 million times cheaper). This huge advance in readily available computer technology is directly due to applied quantum mechanics.



So, *what is quantum mechanics?* An answer to this question requires grappling with the fundamental microscopic schizophrenia of the universe. A good place to begin is to recall some basic ideas from classical mechanics and electromagnetism.

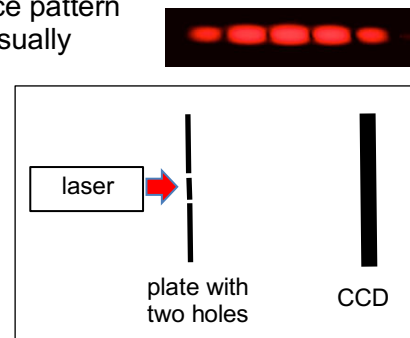
Particles and waves

A “particle” is a classical, Newtonian concept. In classical physics, a particle is usually thought of as a chunk of matter, a localized entity with *ignorable* internal structure—effectively a “point.” (For some purposes, a cow might be treated as a particle if its tail, hooves, and innards are of no particular interest.) A particle is characterized by a small number of **intrinsic** (i.e., independent of motion) properties such as its mass and electric charge. It also carries **extrinsic** (i.e., dependent on motion) properties such as linear momentum, kinetic energy, and angular momentum with respect to some reference point. The unique signature of classical particleness is that the entity's physical properties show up at a “point” in space. A “wave,” on the other hand, is delocalized; the physical properties it carries are spread out in space and only have values when their *densities* are summed over a finite volume. The unique signature of waviness is *interference*, produced when multiple waves overlap.

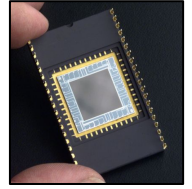
Classical waves: the double slit experiment in bright light

First performed around 1800 by Thomas Young, the interference pattern observed in the double slit experiment (e.g., the figure to the right) is usually

taken as providing convincing evidence that light is a wave phenomenon. The modern version of the double slit experiment involves a laser and a charge-coupled device (CCD) collector (see sketch). Light from the laser illuminates two small slits in an otherwise opaque plate and is eventually absorbed at the CCD collector. A laser is used because its light is “coherent” (which means that light emerging from the holes has a fixed phase



difference) and almost a single color (a single wavelength). The CCD (image to the right) is a semiconductor device consisting of a grid of “pixels”—electrically isolated, capacitive elements that acquire a charge when illuminated with light. CCD collectors are used to record images in digital cameras, for example. The charge on a pixel is collected for a time Δt and is proportional to the light energy absorbed by the pixel during that time. After the collection time, the charges are recorded and the values stored. These, in turn, are used to reconstruct an intensity image with bright (lots of absorbed energy) and dark (not much absorbed energy) regions. This is a single “stationary” interference pattern.



Understanding the double slit experiment: electromagnetic waves

Interference patterns in bright light are explained in terms of overlapping electromagnetic waves emitted from different sources. Electromagnetic waves are electric and magnetic fields that vary in time and space and that “create one another” according to Maxwell’s classical theory. Classical electromagnetism can be summarized succinctly as follows:

- Electric charge makes electric field ($\vec{\mathcal{E}}$)
- Moving electric charge makes magnetic field (\vec{B})
- Time-changing magnetic field also makes electric field
- Time-changing electric field also makes magnetic field

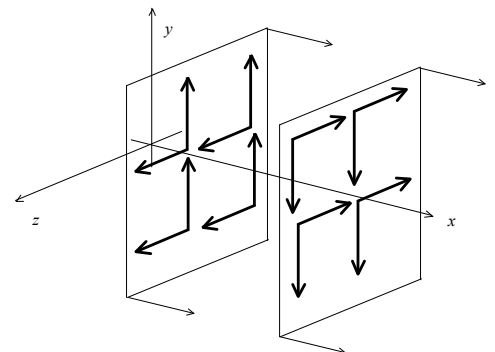
These four points imply that an *accelerating* charge makes a time-changing \vec{B} that, in turn, makes a time-changing $\vec{\mathcal{E}}$, that, in turn, makes a time-changing \vec{B} , that, in turn, makes This sequence of events corresponds to electric and magnetic fields—together called *electromagnetic (EM) radiation*—that propagate away from the accelerating *source* charge. Electromagnetic fields carry energy. Associated with electromagnetic fields at a given point in space at a given time is another field, u , which is the electromagnetic energy per unit volume (the *energy density*):

$$u = \frac{\epsilon_0}{2} (\mathcal{E}^2 + c^2 B^2). \quad (1)$$

In a sense, u is more physically relevant than $\vec{\mathcal{E}}$ or \vec{B} because it directly states where, when, and how much electromagnetic energy is available for interactions with charges.

Maxwell’s equations (MEs) constitute a precise, mathematical statement of how (“bright”) EM radiation originates and sustains itself. MEs predict that the rate of propagation of EM radiation depends on the *electric permittivity* and the *magnetic permeability* of the medium through which it propagates. Vacuum *permits* electric fields ($\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2/\text{Nm}^2$) and can be *permeated* by magnetic fields ($\mu_0 = 4\pi \times 10^{-7} \text{ N/A}^2$), and the speed of propagation in vacuum ($= 1/\sqrt{\epsilon_0 \mu_0}$) equals $c = 3.0 \times 10^8 \text{ m/s}$, the speed of light.

Though hardly the only solution to MEs, the *electromagnetic plane wave* in a vacuum is a particularly useful wave model. The plane wave geometry is depicted to the right. Two parallel, planar sheets of electric and magnetic fields are shown traveling in the positive x -direction. In these sheets, $\vec{\mathcal{E}}$ points in the $\pm y$ -direction and \vec{B} points in the $\pm z$ -direction. In a given sheet, $\vec{\mathcal{E}}$ is the same at every point, and \vec{B} is the same at every point. At one instant, the



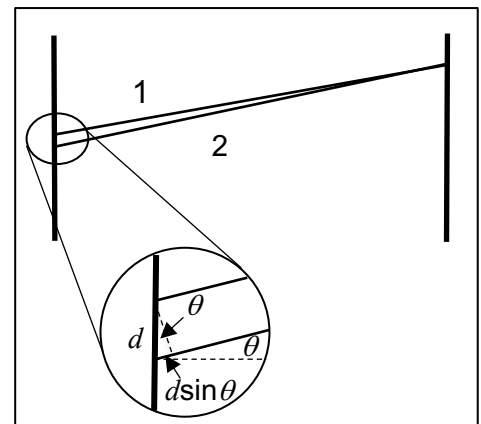
difference in the respective fields is only along the x -direction. **For an EM plane wave in vacuum, the magnitudes of $\vec{\mathcal{E}}$ and \vec{B} are related by $\mathcal{E} = cB$. For a traveling plane wave, B^2 can be eliminated from (1) leaving $u = \epsilon_0 \mathcal{E}^2$; for this case, the energy densities associated with the electric and magnetic fields are equal.** Often, the space and time dependences of the fields are “harmonic”—i.e., can be represented by sines, cosines, or complex exponentials.

Example: $\mathcal{E}(x,t) = \mathcal{E}_{\max} \cos(kx - \omega t + \phi)$ is a sinusoidal electric field wave traveling in the $+x$ -direction. [Why? The quantity $(kx - \omega t + \phi)$ is the *phase* of the cosine function. To travel with the wave is to keep the phase constant. As t increases, x has to increase also so that the phase is constant: the phase is travelling in the $+x$ -direction. If, instead, the phase were $(kx + \omega t + \phi)$, then x would have to decrease to keep the phase constant; such a wave would be traveling in the $-x$ direction.] \mathcal{E}_{\max} is the *amplitude* of the wave; k is the *wavenumber*, $k = \frac{2\pi}{\lambda}$, where λ is the *wavelength*; ω is the *angular frequency*, $\omega = 2\pi f$, where f is the (ordinary) *frequency*; and ϕ is the *phase* of the wave at $x = 0, t = 0$. For example, the electric field amplitude of visible sunlight at the top of Earth’s atmosphere is about 860 V/m, and an average wavelength is about 550 nm—yielding a wavenumber of about 1.1×10^7 rad/m. (**The wavelength range for visible light is from about 400 nm to 700 nm.**) Note that for EM waves $c = \lambda f$; as a result, $kc = 2\pi f = \omega$. The frequency of a 550 nm wavelength EM wave is about 5.5×10^{14} Hz and the angular frequency is 3.4×10^{15} rad/s. Thus, for sunlight, for example, $\mathcal{E}(x,t) = (860 \text{ V/m}) \cos[(1.1 \times 10^7 \text{ rad/m})x - (3.4 \times 10^{15} \text{ rad/s})t + \pi/6]$, where $\pi/6$ is the wave’s (arbitrarily chosen) initial phase.

Example: For the sunlight example above,
 $u(x,t) = (6.55 \times 10^{-6} \text{ J/m}^3) \cos^2[(1.1 \times 10^7 \text{ rad/m})x - (3.4 \times 10^{15} \text{ rad/s})t + \pi/6]$. In other words, **energy propagates at the same speed as the field and has maximum values where the field has its maximum positive and negative values and is zero where the field is zero (its “nodes”).**

EM radiation from the laser used in the double slit experiment causes charges in the plate to accelerate and radiate also. The observed double slit interference pattern results from a superposition of the incident laser radiation with induced radiation from the plate. It is possible to compute the double slit intensity distribution on the collector by just assuming that the holes in the plate each radiate electromagnetic waves of the same wavelength. Though it at first seems unlikely that “holes radiate,” this assumption is justified in Appendix I below.

The geometry of the double slit set-up is depicted to the right. The distance from the plate to the collector is typically very large compared with the slit separation, d , so lines emerging from the slits that join at the collector are almost parallel. The distance from slit 2 to the collector is therefore greater than that from slit 1 by an amount $d \sin \theta$ as shown in the inset. In order for the light along paths 1 and 2 to get to the



collector at the same time, the light from 2 must leave before the light from 1. That means the waves along the two paths will have varying phase differences depending on the angle θ . The total field from the holes is $\mathcal{E}_{\text{holes}} = \mathcal{E}_{\text{max}} [\cos(kx - \omega t + \phi_1) + \cos(kx + kd \sin \theta - \omega t + \phi_2)]$, where the traveling waveform defined in the preceding examples is used. The first cosine is the wave from plug 1 and x is the distance from plug 1 to the collector; the second cosine is from plug 2; the quantities ϕ are initial phases at the two slits. The CCD collects light energy for a time Δt that is many cycles long. Consequently, the total accumulated essentially energy equals the number of cycles times the average accumulated energy per cycle. Thus, the distribution of accumulated energy over the CCD is proportional to $(-\mathcal{E}_{\text{holes}})^2$ averaged over one cycle of the light wave. Appendix II below shows that for laser light the double slit intensity (brightness) interference pattern on the collector is

$$I = 2I_0[1 + \cos(kd \sin \theta + \phi_1 - \phi_2)], \quad (2)$$

where I is the time-averaged absorbed energy and I_0 is the intensity of light from one slit if it were by itself. **If the laser beam is normal to the plate surface**, $\phi_1 = \phi_2$, and in that case I is a maximum ($= 4I_0$) for $kd \sin \theta_{\text{max}} = 2n\pi$ and is zero for $kd \sin \theta_{\text{min}} = (2n+1)\pi$, where in either case n is 0 or a positive or negative integer. Since $k = 2\pi/\lambda$, these two conditions are equivalent to

$$d \sin \theta_{\text{max}} = n\lambda \quad (3)$$

and

$$d \sin \theta_{\text{min}} = [(2n+1)/2]\lambda. \quad (4)$$

Note that irrespective of what d and λ are **there is always a “central maximum”** at $\theta_{\text{max}} = 0$ corresponding to $n = 0$. On the other hand, **there are no other maxima if $\lambda > d$** , because $|\sin \theta| \leq 1$.

Example: Red laser light ($\lambda = 650 \text{ nm}$) illuminates a double slit (head on) with slit separation = 1500 nm. How many maxima can be observed in the interference pattern?

Solution: $\sin \theta_{\text{max}} = n650/1500$. The magnitude of the sine must be no greater than 1, so the magnitude of n has to be 2 or less. Therefore, 5 maxima can be observed, namely, corresponding to $n = 0, \pm 1, \pm 2$.

Example: Suppose the slit separation is now 15000 nm. The collector is 1 m from the plate. How far from the central maximum is the first adjacent maximum (on either side) on the collector?

Solution: $\sin \theta_{\text{max}} = 650/15000 \approx 0.043$, a number much less than 1. For small angles, $\sin \theta \approx \tan \theta$, so the distance from the central maximum is about 0.043 m = 4.3 cm.

Appendix I: Holes “can radiate”

Calculating where the bright and dark spots in the interference pattern will be sounds impossible because we don’t know much about the plate. The one thing we *do* know—namely, that the plate without holes is opaque—saves the day. Here’s how. We want to determine the distribution of $u = \epsilon_0 \mathcal{E}^2$ over the collector. The electric field in this expression is the electric field

anywhere on the collector due to *all* radiating sources. Thus (ignoring that electric field is a vector),

$$\mathcal{E}_{\text{collector}} = \mathcal{E}_{\text{laser}} + \mathcal{E}_{\text{plate w holes}}.$$

But, the plate without holes is opaque so it must be that

$$0 = \mathcal{E}_{\text{laser}} + \mathcal{E}_{\text{plate w/o holes}}.$$

Moreover, we can think of the plate without holes as the plate with *small* holes plus *small* plugs that fit in the holes (this assumes that the radiation from the plugs doesn't alter the radiation from the plate; that's *not* true if the plugs are large); thus,

$$0 = \mathcal{E}_{\text{laser}} + \mathcal{E}_{\text{plate w holes}} + \mathcal{E}_{\text{plugs}},$$

or what is the same thing

$$\mathcal{E}_{\text{collector}} = -\mathcal{E}_{\text{plugs}}.$$

The left hand side of this equation is what we want. We can calculate it from the right hand side by treating the holes as little radiating plugs. We don't need to know anything about the plate beyond that it is opaque when its holes are filled. Cool.

Appendix II: More about double slit intensities

For notational simplicity, let $\alpha = kx - \omega t + \phi_1$ and $\beta = kx + kd \sin \theta - \omega t + \phi_2$. Then the desired energy distribution is proportional to

$$\langle (\cos \alpha + \cos \beta)^2 \rangle = \langle \cos^2 \alpha \rangle + \langle \cos^2 \beta \rangle + \langle 2 \cos \alpha \cos \beta \rangle,$$

where the brackets mean "average over 1 cycle." Both α and β depend linearly on time so the average of the respective cosines is zero and the average of the respective cosine²s is $\frac{1}{2}$. Thus, the sum of the first two terms on the right hand side is 1. The value of the third term isn't immediately obvious because α and β are shifted in phase by $kd \sin \theta$. A trig identity comes to the rescue: $2 \cos \alpha \cos \beta = \cos(\alpha + \beta) + \cos(\alpha - \beta)$; note that

$$\alpha + \beta = 2kx - 2\omega t + kd \sin \theta + \phi_1 - \phi_2$$

while

$$\alpha - \beta = -(kd \sin \theta + \phi_1 - \phi_2).$$

Because $\alpha + \beta$ depends linearly on time

$$\langle \cos(\alpha + \beta) \rangle = 0.$$

There is no explicit time dependence in $\alpha - \beta$; if $\phi_1 - \phi_2$ does not change in time then

$$\langle \cos(\alpha - \beta) \rangle = \cos(kd \sin \theta + \phi_1 - \phi_2).$$

Extended light sources other than lasers are *incoherent*, which means that $\phi_1 - \phi_2$ fluctuates randomly in time. In that case, $\langle \cos(\alpha - \beta) \rangle = 0$; consequently, the interference term in (2) is zero. This is what a laser does for you: laser light is *coherent* and keeps $\phi_1 - \phi_2$ fixed in time.

Note that the maximum value of I in (2) ($= 4I_0$) is *two times greater* than the sum of the maxima of the two slits individually. Of course, the minimum value is zero, so the average over the collector comes out to be just the sum of the intensities from the two slits separately. For incoherent light the intensity on the collector is uniform (we've ignored the fact that the electric field decreases with distance from the source) and is just the sum of the intensities of the two slits.