What is a photon?

Dan Piponi

August 9, 2017

Abstract

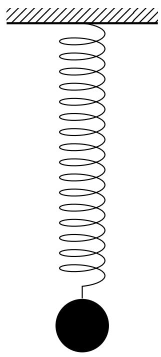
Popular science writing about quantum mechanics leaves many people full of questions about the status of photons. Unfortunately, complete answers to many of these questions require detailed calculations using techniques from quantum field theory. Nonetheless, there are still many questions that can be answered without tricky mathematics.

One of the challenges is that photons are very different to ordinary everyday objects like billiard balls. This is partly because photons are described by quantum mechanics whereas billiard balls are better modelled with classical Newtonian mechanics. Quantum mechanics defies many of our intuitions. But it's also because the word *photon* plays by different linguistic rules to *billiard ball*. I hope to explain why in a way that uses a minimum of mathematics, and introduce some interesting physics along the way.

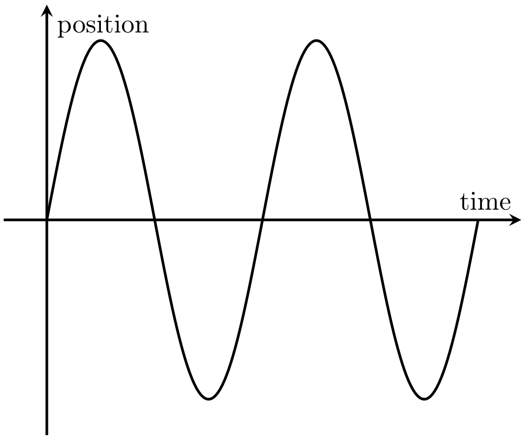
One of my goals is to avoid saying anything original. I'm largely going to recapitulate, minus the mathematics, material I first learnt from three or so courses I took at Cambridge University many years ago: Quantum Mechanics, Solid State Physics and Quantum Field Theory. I also learnt about some of this from David Miller at Stanford University who talked a little about what properties it is meaningful to apply to a photon. (I hope I haven't misrepresented him too badly.)

1 The simple harmonic oscillator

Here's a mass hanging on a spring:



Suppose it's initially sitting in equilibrium so that the net force acting on it is zero. Now we lift the mass a small distance and let it go. Because we lifted it a bit, we shortened the spring, reducing the tension in the spring. This means the force due to gravity is now more than the spring tension and the mass falls. Eventually it balls below the equilibrium point, increasing the tension in the spring so it's more than the force due to gravity. The result is a net force that pulls the mass back up again. To a good approximation, the force restoring the mass to its equilibrium point is proportional to how far it has been displaced. When this happens we end up with oscillating motion where the mass bounces up and down. Here's what a graph of its displacement looks like over time:

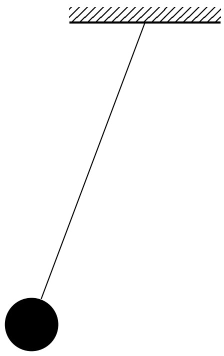


It's actually a sine wave but that detail doesn't matter for us right now.

An oscillator where the restoring force is proportional to the displacement from the equilibrium point is called a *simple harmonic oscillator*[1] and its oscillation is always described by a sine wave. I assume the sine wave is what gives it its name: this is the kind of oscillation you get from a musical instrument playing a pure note at one frequency.

Note that I'm ignoring friction here. This is a reasonable approximation for many physical systems.

Masses on springs aren't all that important in themselves. But many other physical systems also involve an oscillation about some equilibrium point with a restoring force proportional to the displacement. In fact, it's not completely unreasonable to say that almost all systems involving a single particle in some kind of force field result in simple harmonic motion when the displacements from equilibrium are small enough. Another standard example is the pendulum swinging under the influence of gravity:



At a more fundamental level, an example might be an atom in a crystal being held in place by electrostatic forces from its neighbouring atoms.

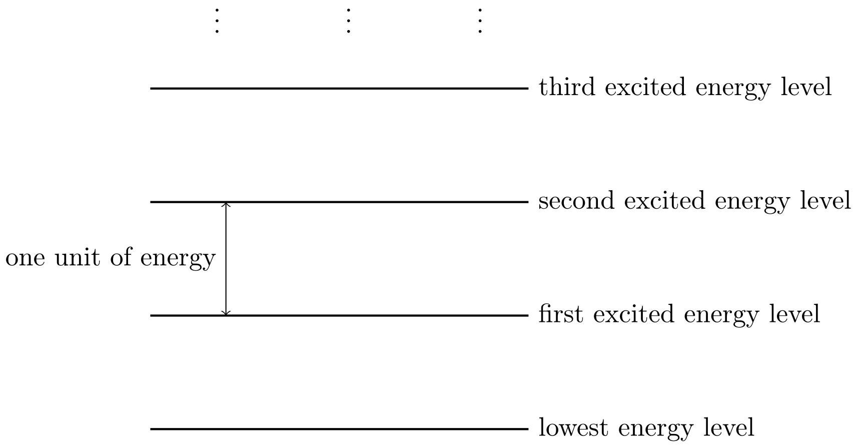
If you have one of these systems, then in principle you can set it in motion with as little energy as you like. Pull a mass on a spring down a little bit and it will bounce back up, oscillating a certain amount. Pull the mass down half the amount and it'll bounce with oscillations half the size. In principle we could keep repeating this experiment, each time starting with the mass displaced half the amount we tried previously. In other words, a simple harmonic oscillator can have any energy we like. The *spectrum* of possible energies of one of these oscillators is continuous. (Note that the word *spectrum* here is merely physicist-speak for a set of possible values.) If we can set one in motion with 1 unit of energy then we can also set it oscillating with 0.5 units, or 0.01 units, or 0.000123 units of energy.

2 Quantum mechanics

Everything I've said above is assuming that classical Newtonian mechanics is valid. But we know that for very small systems, around the size of a few atoms or smaller, we need to use quantum

[1] See `https://en.m.wikipedia.org/wiki/Simple_harmonic_motion`

mechanics. This is an enormous topic but I'm only going to extract one basic fact. According to quantum mechanics, a simple harmonic oscillator isn't free to oscillate with any energy you like. The possible energy levels, the spectrum of the system, is discrete. There is a lowest energy level, and then all of the energy levels above that are equally spaced like so, going up forever:



We usually call the lowest energy level the *ground state*[2] or *vacuum state* and call the higher levels *excited* states.

The spacing of the energy levels depends on the *stiffness* of the system, which is just a measure of how much the restoring force increases with displacement from equilibrium. Stiffer systems will have a higher frequency of oscillation and a bigger spacing between the energy levels.

(I'm deliberately not saying anything about why we get discrete energy levels in quantum mechanics. I just want to use this one fact so I can get on and talk about photons eventually.)

In practice the difference in energy between one level and the next is tiny. This means that if you're literally fiddling about with a mass on a spring you won't ever feel the discreteness. The amount your hand trembles is many orders of magnitude greater than the effect of this discreteness. Nonetheless, it is extremely important when modeling microscopic systems.

3 Quantum linguistics

Suppose you have two identical simple harmonic oscillators next near each each other. They each start at some energy level. Then let's say they interact in some way that I'm not going to specify. Now suppose the interaction is eventually switched off. They'll each end up at some energy level again. Because of energy conservation the total energy will be the same but one of the oscillators might end up at a higher energy level while the other ends up at a lower energy level.

Here are some English sentences we could say about the kinds of systems I've described so far:

1. This system is in the ground state.
2. That system is in its first excited state
3. This system is at an energy level higher than that system
4. During the interaction the energy level of this oscillator went down and the energy level of that one went up

Now I want to introduce the (count) noun *quantum*[3], with plural *quanta*. The idea here is not that I'm telling you about a new entity. I want to present this as a new way to talk about things I've already introduced. So rather than give a definition of *quantum* I will instead show how you can rewrite the above sentences using the language of quanta:

1. There are no quanta in this system

[2] See `https://en.m.wikipedia.org/wiki/Vacuum_state`
[3] See `https://en.m.wikipedia.org/wiki/Quantum`

2. That system has one quantum of energy
3. This system has more quanta than that one
4. Some quanta were transferred from this system to that system.

Those sentences make it seem like I'm talking about a new kind of object - the quantum. But I'm not. They're just a manner of speaking about energy levels. I hope I've given you enough examples to get the idea.

Just in case you think it's weird to talk about energy levels in terms of quanta, I'd like to remind you that you already do this all the time with money. Dollar bills are actual objects that exist in the world. But money in your bank account isn't. Somewhere in some database is a representation of how much money you have. You might say "I have one hundred dollars in my savings account" But those dollars certainly don't exist as distinct entities. It doesn't really make sense to talk about the thirty-seventh dollar in your bank account. You can transfer dollars from one account to another, and yet what's really happening is that two totals are being adjusted. We treat these accounts a lot like they're containers holding individual objects called dollars. Certainly our language is set up like that. But we know that it's really just the totals that have any kind of representation. The same goes for quanta. It's just a manner of speaking about systems that can have different amounts of energy and where the spectrum of energy levels forms a ladder with equally spaced rungs. Because of your experience with money I probably don't need to give you any more examples.

One more bit of terminology: when the spectrum of energies is discrete it's said to be *quantised*.

4 Real consequences

This discussion about whether or not quanta have their own individual identities sounds like philosophical nit-picking. But it has real physical consequences.

In thermodynamics we usually assume that, in the absence of other knowledge, all of the possible states of a physical system are equally likely. For example, the behaviour of gases can be understood as a consequence of assuming that all arrangements of molecules in a container are equally likely, subject to things we might know about the gas like its total mass and energy.

So let's think about what must be almost the simplest thermodynamic system possible: two quantum harmonic oscillators with the same frequency and where we know that there are two quanta shared between them.

If the quanta are genuinely distinct entities there are four possibilities:

1. Both quanta are in the first oscillator.
2. The first quantum is in the first oscillator and the second one is in the second oscillator.
3. The first quantum is in the second oscillator and the second one is in the first oscillator.
4. Both quanta are in the second oscillator.

Half the time the quanta end up in the same oscillator.

Now suppose the quanta aren't distinct entities but the complete state of a harmonic oscillator is given by its energy level. Then we have just three possibilities:

1. Both quanta are in the first oscillator.
2. There is one quantum in each oscillator.
3. Both quanta are in the second oscillator.

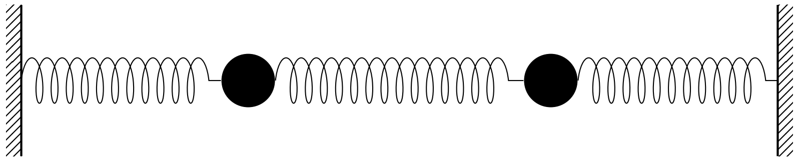
Now both quanta end up in the same oscillator two out of three times.

What happens in reality? The latter case. Indistinguishable quanta are more likely to hang out together than distinguishable ones. In fact, there is a name for quanta that do this: they're called *bosons*[4]. An extreme example of this is the Bose-Einstein condensate[5] where it's possible to get a large number of bosons to enter the same state. This is observable in the lab. Boson statistics also determines the colour of the sun but I won't get as far as explaining that.[6]

[4] See `https://en.m.wikipedia.org/wiki/Boson`
[5] See `https://en.m.wikipedia.org/wiki/Bose-Einstein_condensate`
[6] See `https://en.m.wikipedia.org/wiki/Planck's_law`

5 Coupled systems

Let's return to classical physics with a slightly more complex system consisting of two masses connected to springs. We ignore gravity now:



We restrict ourselves to just considering back and forth motion constrained along a horizontal line. This is a coupled system. If the left mass moves to the right, not only does it experience a restoring force pushing it left, but the mass on the right will experience more of a force pushing it to the left. We can't treat the masses as independent and so we don't get the simple solution of each mass always oscillating with a sine wave.

For this particular problem though there is a trick to turn it into a pair of harmonic oscillators. The idea is to consider the pair of masses as a single entity. We can think of the centre of mass of the pair, the midpoint between them, as being one variable that describes this entity. Let's call its motion the *external* motion. We can also think of the distance between the two masses in the pair as being the system's *internal* motion. (I'm just using *internal* and *external* as convenient names. Don't read too much into them.) It turns out that when you analyse this using classical dynamics the internal motion and the external motion act like independent quantities. What's more, each one behaves exactly like it's simple harmonic. So we get one sine wave describing the overall motion of the pair, and another one that describes how the elements of the pair oscillate with respect to each other. In a way that's similar to the case with one mass on a spring we can initialise the motion with whatever energy we like. Except in this case we can choose any energy we like for the internal motion and any energy we like for the external motion. (Up to a point. Too much energy and the masses start whacking the walls on the side. So we're more interested in the low energy case here.)

Interestingly the frequencies of the internal and external motions are typically different. So you can end up with some quite complicated motions with two different frequencies beating against each other.

When we're able to find ways to split up the motion into independent quantities, each of which is simple harmonic, each kind of motion is said to be a *normal mode*[7].

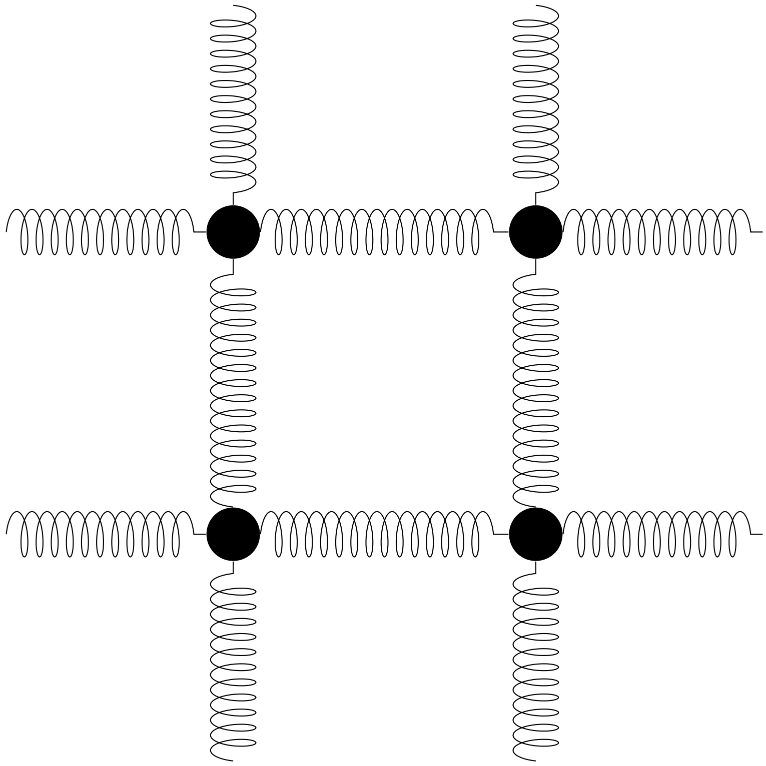
When you have independent normal modes, the theory behind quantum mechanics says we can treat them independently. So what we expect is that the spectrum of possible energy levels for this system is, in some sense, two-dimensional. We can put quanta into the internal oscillation and we can put quanta into the external oscillation. Note that these quanta are associated with different frequencies of oscillation so they correspond to different amounts of energy.

(And a reminder: when talking about quantum mechanics I'm not talking literally about masses on springs. I'm talking about physical systems that have equations of motion that mean they behave like masses on springs. In this case it might be a pair of particles trapped in a microscopic well with a repulsive force between them.)

6 Solid state physics

Now I'm going to jump from just two masses to a large number of them. For example, the behavior of trillions of atoms in a crystal can be approximately modelled by a grid of masses and springs, of which the following diagram is just a tiny piece:

[7] See https://en.m.wikipedia.org/wiki/Normal_mode



A real crystal would be arranged in a 3D lattice but I've drawn 2D here for convenience.

Think of the springs as both pushing apart atoms that get close, and pulling together atoms that move apart.

This is a highly coupled system. Ultimately every atom in our lattice is connected to every other one, either directly, or indirectly. Nonetheless, it is still possible to find normal modes. The normal modes all have the same basic form: for each normal mode there is some direction, velocity and frequency, and the normal mode consists of sinusoidal waves of displacement travelling in the given direction with the given speed and oscillation frequency. These waves extend through the entire crystal, but with fixed spacing between planar wavefronts. These waves are known as plane waves. If the system is perfectly harmonic, so the restoring force is precisely proportional to the displacement, then each direction and frequency of wave oscillates its way through the crystal completely independently of any other. Just as how in the example with two masses any possible oscillation is a combination of internal and external motion, for a crystal lattice any motion is a combination of these plane waves. (Decomposing any oscillation as a combination of plane waves is known as computing its Fourier transform.[8])

Now we're ready to consider this situation quantum mechanically. Because each plane wave is a normal mode, we can treat each one as an independent simple harmonic oscillator. This means that the energy in each plane wave is quantised. So when we consider a crystal lattice quantum mechanically we find that its states consist of plane waves propagating through it, but where the amount of energy in each wave is given by a discrete spectrum. So again we can talk about how many quanta there are in each mode.

Linguistically it gets a bit more interesting now. Each plane wave is associated with a particular direction and speed so it makes sense to talk of these quanta as have a direction and speed. But note that statements involving quanta are still really just sentences about energy levels. So, for example, the statement "the mode of this system with this velocity and frequency is in its first excited state" is, by definition, exactly the same as "this system has precisely one quantum with this velocity and frequency". In particular, when we write sentences like these we aren't implying that there is some new kind of object, the quantum, that has suddenly attached itself to our crystal. The quanta are, in effect, properties of the lattice. By the way, in the particular case of vibrating atoms in a lattice, the quanta are known by a special name: *phonons*[9].

[8]See `https://en.m.wikipedia.org/wiki/Fourier_transform`

[9]See `https://en.m.wikipedia.org/wiki/Phonon`

7 Quantum field theory and photons

And now we're ready to move onto photons.

In classical physics, electromagnetism is described by Maxwell's equations. Maxwell's equations say that a varying magnetic field generates an electric field and a varying electric field generates a magnetic field. The result is that it is possible for an oscillating electric field to create an oscillating electric field so that an electric field can propagate through space on its own without the help of electric charges or electric currents or any other kind of 'generator'. As these electric fields also produce magnetic fields that propagate with them, the whole thing is called an electromagnetic wave.

Just like in the case of displacements in a crystal lattice, an electromagnetic wave also has normal modes. For each choice of direction and frequency, plane waves with those parameters form a normal mode. In fact, you have personal experience of this fact. Visible light is electromagnetic radiation with a frequency of around 500 THz. Wifi uses signals at around 5 GHz. The radio might use signals at around 100 MHz. When you surf the web wirelessly while listening to the radio, the wifi signals don't interfere with your vision or the radio signal. (Actually, wifi might interfere with the radio signals, but not because of the 5 GHz signals. It might happen if badly manufactured hardware emits stray signals around the 100 MHz band.) That's because these waves pass through each other without being coupled to each other in any way. And at this point you must already be guessing what a *photon* is. For each choice of frequency and direction (and also polarisation, but that's just a detail) the amount of energy that can be in the corresponding mode is quantised. For the electromagnetic field the quanta are called photons.[10]

And that's it! Electromagnetic waves can be thought of as being made up of different oscillation modes. Because of quantum mechanics, each mode contains an amount of energy that is quantised to be a whole number multiple of some base amount. Although the thing that really matters is the total amount of energy in the mode (like how it's your total bank balance that matters) it can still be useful to talk about this total as if it's a collection of objects called photons, even though there isn't really an object corresponding to the noun (and just as there aren't really objects corresponding to the individual dollars in your account).

8 Some consequences

The normal modes for an electromagnetic wave are plane waves that are extended in space. In principle all the way across the universe but for practical problems physicists often consider electromagnetic waves in a large but finite box. This means that adding a quantum to a system has an effect that extends across the entire system. That makes it highly problematic to talk about the location of a photon. The notion doesn't really make any sense.

In classical electromagnetism the normal modes of an electromagnetic field carry not only energy but also momentum. So it makes sense to talk of the momentum of a photon. Again, I must reiterate that we talk of the momentum of the photon but really we're talking about a property of the field.

Note that the last two paragraphs are consistent with the Heisenberg uncertainty principle. In this particular case it says that if we know the exact momentum of a photon then we must be completely uncertain about its position. That's exactly the situation we have. Photons correspond to normal modes with a precisely specified energy, frequency and momentum. Additionally they are associated with plane waves that extend across space so we can't specify their position.

9 Stuff I left out

I just wanted to answer one question. To do this I had to leave out a lot of stuff:

1. Fermions.
2. Virtual photons and Feynman diagrams.
3. Superpositions.

[10] See `https://en.m.wikipedia.org/wiki/Photon`

4. Quasiparticles.
5. The way light beams at a low enough intensity look like streams of discrete particles.
6. Wave/particle duality.
7. Feynman's description of photons in his book QED.

All I can do is apologise.