

Photons and Polarization

Now that we've understood the classical picture of polarized light, it will be very enlightening to think about what is going on with the individual photons in some polarization experiments. This simple example will reveal many features common to all quantum mechanical systems.

Classical polarizer experiments

Let's imagine that we have a beam of light polarized in the vertical direction, which could be produced by passing an unpolarized beam of light through a vertically oriented polarizer. If we now place a second vertically oriented polarizer in the path of the beam, we know that all the light should pass through (assuming reflection is negligible). On the other hand, if the second polarizer is oriented at 90 degrees to the first, none of the light will pass through. Finally, if the second polarizer is oriented at 45 degrees to the first (or some general angle θ) then the intensity of the transmitted beam will be half (or generally $\cos^2(\theta)$ times) the intensity of the original polarized beam. If we insert a third polarizer with the same orientation as the second polarizer, there is no further reduction in intensity, so the reduced intensity beam is completely polarized in the direction of the second polarizer.

Photon interpretation of polarization experiments

The results we have mentioned do not depend on what the original intensity of the light is. In particular, if we turn down the intensity so much that the individual photons are observable (e.g. with a photomultiplier), we get the same results.¹ We must then conclude that each photon carries information about the polarization of the light. For example, in the first two cases above, all of the photons pass through when the second polarizer is aligned with the first, but none of the photons pass through if the second polarizer is perpendicular to the first. So each photon must "know" that it is part of a vertically polarized light beam. In other words, it must be that polarization is a property of the individual photons, rather than a property (like intensity) that requires us to know about many photons at once.

Now let's try to understand what is going on when the second polarizer is oriented at 45 degrees to the first one. In this case, the transmitted intensity

¹In the case where the intensity is so low that the photons arrive sporadically, intensity is defined as the average over a long period of time of the energy per unit time.

is half the intensity of the original polarized light. In terms of the photon picture, intensity is proportional to the number of photons passing in a given amount of time times the energy of each photon. But since the polarizers do not change the wavelength of light passing through, the energy of the photons passing through the polarizer are the same as the ones incident on the polarizer. So the only way to explain the 50 percent reduction in intensity is to say that only half of the vertically polarized photons pass through the polarizer oriented at 45 degrees. Furthermore, the photons that do pass through will all pass another polarizer at 45 degrees, so the polarization state of the original photons is also changed to 45 degree polarization.

Our conclusion sounds innocent enough, but it is dramatically different what we are used to in classical physics. What we are saying is that with a stream of identical vertically polarized photons, half of the photons pass through the polarizer and half do not. Further investigation reveals that the distribution of photons that pass through and photons that are absorbed is random in time, just like the results of a series of coin flips. So we can say that each photon has a 50 percent probability of passing through the polarizer and a 50 percent probability of being absorbed. It is only these probabilities, and not the actual results of the experiment that are determined by the initial conditions.²

It is easy to generalize our conclusions to the case where the second polarizer is at some angle θ to the first one. In this case, the transmitted intensity is $\cos^2(\theta)$ times the incident intensity. To reproduce this using the photon picture, we must assume that the photons are transmitted with a probability $\cos^2(\theta)$ and absorbed with a probability $\sin^2(\theta)$.

Mathematical model for classical polarization experiments

How can we come up with a model to explain these results? Let's first go back and think about the classical explanation of why the intensity is reduced. For definiteness, let's say we have light propagating in the \hat{z} direction, so the allowed polarizations are in the plane formed by \hat{x} and \hat{y} . We'll assume that we have a polarizer oriented so that light polarized in the \hat{x} direction (call this vertical) passes through and light oriented in the \hat{y} direction (call this

²Note that in the coin flip example, knowing the precise initial configuration and velocity of the coin would allow us to predict whether we'll get heads or tails. For the photons, we are claiming that the results could be different even if the initial conditions are identical.

horizontal) is absorbed.

Now, in the classical picture, light with some general polarization may be written as a SUPERPOSITION of light polarized parallel to the polarizer and light polarized perpendicular to the polarizer. For example, if we have a wave with amplitude E_0 polarized at an angle θ to the vertical direction, this is a linear combination of a vertically polarized wave with amplitude $E_0 \cos(\theta)$ and a horizontally polarized wave with amplitude $E_0 \sin(\theta)$. In terms of the maximum electric field vector, we have

$$\vec{E}_0 = E_0 \cos(\theta)\hat{x} + E_0 \sin(\theta)\hat{y}.$$

Now, if we think about these two components as two separate waves that are superposed on one another, the first will pass through the polarizer while the second will be absorbed. The transmitted wave will be a vertically polarized wave with amplitude $E_0 \cos(\theta)$, with maximum electric field vector

$$\vec{E}_0^t = E_0 \cos(\theta)\hat{x}.$$

Since intensity is proportional to the square of amplitude, we conclude that the fraction of light transmitted is the square of transmitted amplitude over incident amplitude, or equivalently, the squared length of the vertical component of \vec{E}_0 over the squared length of the \vec{E}_0 . This gives $\cos^2(\theta)$, thus explaining the observed reduction in intensity.

Mathematical model for photon transmission probabilities

Armed with this understanding, let's now try to come up with a mathematical model for calculating the probabilities of photons being transmitted. Since the results for the probabilities must be the same as the classical results for the intensity fraction transmitted, we should be able to use essentially the same mathematical model, but with a different interpretation. That is, for a given photon polarization state, we will associate a vector \vec{e} pointing in the direction of the polarization. We can think of this as the electric field vector for an individual photon. Then the probability that the photon passes through a polarizer will be the ratio of the squared length of the component of \vec{e} in the direction of the polarizer and the squared length of \vec{e} .

Now, one difference from the classical picture is that all the photon polarization states have the same energy,³ so we can assume that all the vectors

³Here, we are just talking about different polarization states of photons of the same wavelength.

representing individual photons have the same length. If we choose this length to be 1, so that \vec{e} is a UNIT VECTOR, then the probability for transmission is simply equal to the squared length of the component of \vec{e} along the polarizer direction.

A second difference from the classical picture is that when a photon is transmitted, it is not a fraction of a photon, but rather the whole photon. So the vector representing the final state is not just the component of \vec{e} along the polarizer direction (which has length less than one) but a unit vector in the polarizer direction.

Useful notation for unit vectors

To write everything down mathematically, it is convenient to introduce a notation for unit vectors where $|\theta\rangle$ is the unit vector at an angle θ from the \hat{x} direction towards the \hat{y} direction. So we have

$$\hat{x} \equiv |0^\circ\rangle \quad \hat{y} \equiv |90^\circ\rangle \quad |\theta\rangle = \cos(\theta)|0^\circ\rangle + \sin(\theta)|90^\circ\rangle$$

Essential features of the model: the rudiments of quantum mechanics

The model we have come up for calculating transmission probabilities was a simple reworking of the classical model for calculating transmitted intensity. However, it turns out that this model displays most of the basic principles of quantum mechanics. In the summary below, we highlight the essential features that will turn out to be common to all quantum systems.

- **The states** (in this case, the polarization states of photons) **are represented by unit vectors.**
- **For a given experiment or measurement,**(in this case, sending photons towards a polarizer with a fixed orientation) **there are special states whose fate is completely determined. These are called the EIGENSTATES.** For a given polarizer, the eigenstates are the photons polarized in the direction of the polarizer (which definitely pass through), and those polarized in the perpendicular direction (which are definitely absorbed). For example, for a polarizer oriented at 0 degrees (i.e. vertically), the eigenstates are $|0^\circ\rangle$ and $|90^\circ\rangle$.

- **Any general state can be written as a linear superposition of the eigenstates.** Mathematically this is the statement that the eigenstates form a basis for all possible vectors. For example, for a photon polarized in the direction 30 degrees from vertical, we have

$$|30^\circ\rangle = \frac{\sqrt{3}}{2}|0^\circ\rangle + \frac{1}{2}|90^\circ\rangle$$

- **When the experiment or measurement takes place**(in this case, when the photons hit the polarizer) **the state always behaves just like one of the eigenstates. The probability that it behaves like a given eigenstate is the squared length of the unit vector's component along the corresponding eigenstate unit vector.** For the photons, this means that a general photon state either completely passes through and appears as a photon polarized in the polarizer direction, or it is completely absorbed. For example, from the previous equation, the photon in state $|30^\circ\rangle$ will have a 3/4 probability of being transmitted and a 1/4 probability of being absorbed.
- **The state is generally changed by the experiment/measurement.** That is, unless the photon is initially an eigenstate, the state of the photon is changed by the polarizer to become one of the eigenstates.

We'll see that a common feature in all quantum experiments is that what we can predict is not the definite outcome of an experiment, but only the probabilities for various possible outcomes. In our model, the fact that states are represented by unit vectors and probabilities by the squared components guarantees that the probabilities add up to 1. This simple idea turns out to be the right mathematical framework for calculating probabilities in all quantum systems, and this is the reason why so many of the features here hold generally.

In the general case, the unit vectors representing states of quantum systems are not directly related to some classical vectors as they are here, and in fact there are typically more than two (and often an infinite number) of eigenstates.