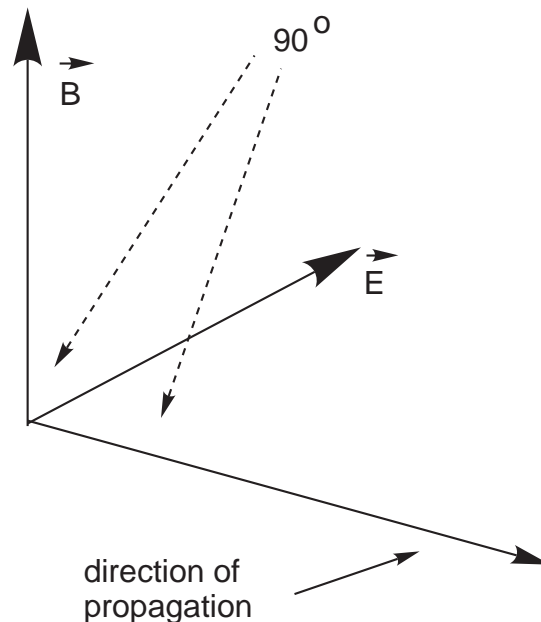


Polarization

Introduction

We will make a first pass at photon polarization in this section and return to a complete theory later on. Here we will introduce polarization and some of its properties and then discuss some experiments that illustrate the problems with the classical description and the strangeness of quantum phenomena.

Light is really oscillating \vec{E} (electric) and \vec{B} (magnetic) field vectors. In particular, these vectors oscillate in time in a plane(two-dimensional space) perpendicular to the direction of propagation (motion) of the light (called the transverse direction). This is why you are able to see light. Your eye is actually responding to these oscillating electromagnetic fields. The oscillations are of very-high frequency (10^{15} /sec)



As stated, the electric field vector is oscillating in some direction at any instant of time. Suppose we choose to associate that **instantaneous direction** of the \vec{E} vector with a new quantity called **polarization**.

The \vec{E} -vector then defines the **plane** of polarization(orthogonal to direction of propagation) and the **direction** of polarization of the beam of light.

Have you ever observed polarization?

Yes...you all have oriented the aerial of portable radios, FM radios or TV rabbit ears to get the best signal...you were lining them up with the polarization of the electromagnetic waves in each case.

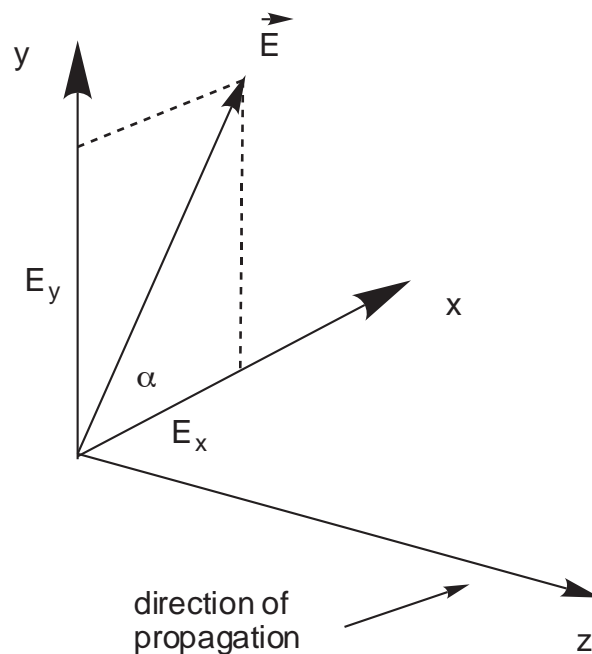
Also, as we shall investigate in much detail shortly, you all have used polaroid sunglasses. These glasses only allow light with

polarization in a single direction to get through, thereby cutting down the intensity(amount) of the light reaching your eyes.

Now for some live experimental demo (watch these carefully...). They contain more than it seems at first sight and actually contains everything we need to know about quantum mechanics.... **Stop by my laboratory (Science Center 127) and try them out for yourselves.**

Now we can **always** choose the plane containing the electric and magnetic fields to be the x - y plane and represent the \vec{E} vector as the **sum** of a two vectors, one in the x -direction and the other in the y -direction, such that the \vec{E} vector makes an angle α with the x -axis.

Then we say that the \vec{E} vector is a combination(sum) of an \vec{E}_x -vector(field) and an \vec{E}_y -vector(field) as shown below.



In terms of the polarization, we say that the direction of polarization as represented by the \vec{E} vector is a (linear) combination of two polarizations, one in the x -direction and one in the y -direction, i.e., the polarization **state of a system** is a linear combination of an x -polarized state and a y -polarized state.

Shortly, we will see how to calculate how much of each state is present in the total state.

Many materials have a property that they will only let light pass through them if the polarization is in a particular direction, i.e., it is in a particular or definite polarization state. Call this axis the **special or preferred or optic axis**. To get through the light must be polarized (have direction of polarization) in a direction **parallel** to the special axis. As I said earlier, you have all seen such material in sunglasses or sheets of **Polaroid** film.

Let us look at some experiments involving polaroid materials, a block of calcite and a laser.

The laser we use produces **unpolarized light** -- what does that mean?

It means that if we check(measure) the amount of polarization in any direction(amount of light getting through an oriented polaroid), then we will find the **same amount** of polarization in all directions.

Experimental Checks:

(1)

Send the laser beam into a polaroid. Our only observation is that the brightness or intensity decreases (if we carefully measured it...it would drop by about 1/2). If we **rotate the polaroid we observe no change**, which says that the amount of light in the beam with polarization parallel to the preferred direction is the same no matter how I orient the polaroid(how I point the preferred direction).

The fact that we get the same intensity (1/2 of total) no matter what direction we choose.....says that the light coming out of the laser has equal amounts of polarization in the two directions(no matter which you two you choose). This is called **unpolarized light** (we will give a more formal definition later).

So,if we choose a preferred direction(done when I pick up the polaroid and orient it in the apparatus), then the polarization of any system(the light from the laser in this case) we are investigating must be thought of as either be parallel or perpendicular to that chosen direction....

no other cases will occur in the real world

After passing through the polaroid we say that the light is polarized in a particular direction (parallel to the polaroid preferred axis). All the light with polarization in the orthogonal direction has been removed(that is what the polaroid does).

If this is true, then if I use 2 identical polaroids rotated by 90 degrees with respect to each other, how much light will come out?

The answer must be none!

So we add another polaroid and rotate so that no light comes through. These two polaroids are now oriented at right angles with respect to each other and since there are only these two possible polarizations with respect to the polaroid directions, no light comes out.

If we rotate the polaroids, keeping there relative orientation fixed no light comes through. This means that there are only two polarizations with respect to the new orientations also and they have cancelled out also.

Again this says that the laser is emitting unpolarized light.

In this experiment, we choose a preferred direction by bringing the

polaroid over. At that point, the light can be thought of as being partly made up of light polarized parallel to the preferred direction and partly of light polarized perpendicular to the preferred direction and the truly amazing result, as we shall see in our later discussions, will be that it "**doesn't decide**" until I bring over the polaroid.

Think about that statement carefully! It reflects the strange sounding explanations that result when we try to use "words" to describe what is happening. Mathematically, as we shall see later, there will be no confusion about what is happening.

(2)

Now we send the laser beam into a calcite crystal and we get two beams. Calcite also has an intrinsic(**built-in**) preferred direction. Calcite causes bending of a light beam with the amount of bending depends on whether the polarization is parallel or perpendicular to the calcite's preferred direction. Since there are only two possible polarizations with respect to the preferred direction of the calcite crystal, we now see two beams.

If I rotate the calcite what will I find?

The same answer with respect to the new preferred direction, of course, i.e., just two beams.

Thus, "the quantum world" is not set until I walk over and make a decision and fix the orientation of the calcite crystal.

Again, think about that statement for a carefully!!! Reflect on the strange explanation in these words.

The physical world is waiting for the observer to make a decision about what measurement will be made. This strange idea, which makes no sense in the classical world, is going to get us into a lot of very hot water later in our discussions.

This result is connected to the fact(as we shall see later) that the act of **measuring** or **finding out** some property of a system in some way determines the state of the system being measured.

Now back to the calcite beams.

We check with a polaroid to see that each of the two beams represents light that has a definite polarization by rotating the polaroid.

Using 2 polaroids, where the 1st cuts out some light and the 2nd cuts out some more, if the rule above is correct and we orient them so that their preferred axes are at right angles, then all light should be stopped. They affect different beams indicating that the beams represent different (orthogonal polarizations).

Mirrors have no effect on polarization....they just redirects beams.

(3)

Now remove the calcite crystal and rotate the pair of polaroids until

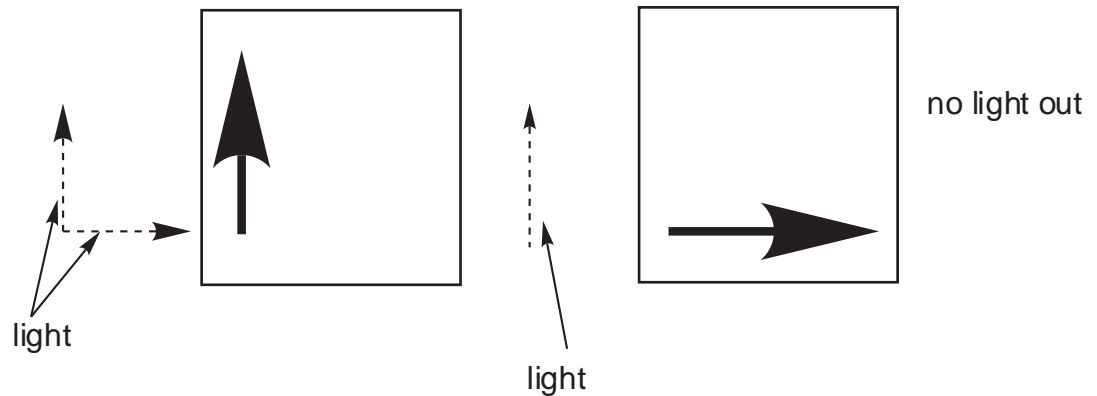
there is no light intensity. Then rotate the second polaroid 45 degrees. The light intensity returns (actually 1/2 of the 0 degree intensity).

Let us see how classical physics explains these polaroid experiments for **intense** (many photon) beams of light.

Classically, if an incident beam of light is polarized parallel to the special or optic axis, then it **all of its energy** gets through the polaroid. If it is polarized orthogonal to the optic axis, then none of its energy will get through the polaroid. If it is polarized at an angle α to the preferred axis, then a fraction $\cos^2\alpha$ of its energy gets through the polaroid

Therefore, we can explain polarization experiments when the notion of electric field vectors make sense (when we have intense light or lots of photons) as follows.

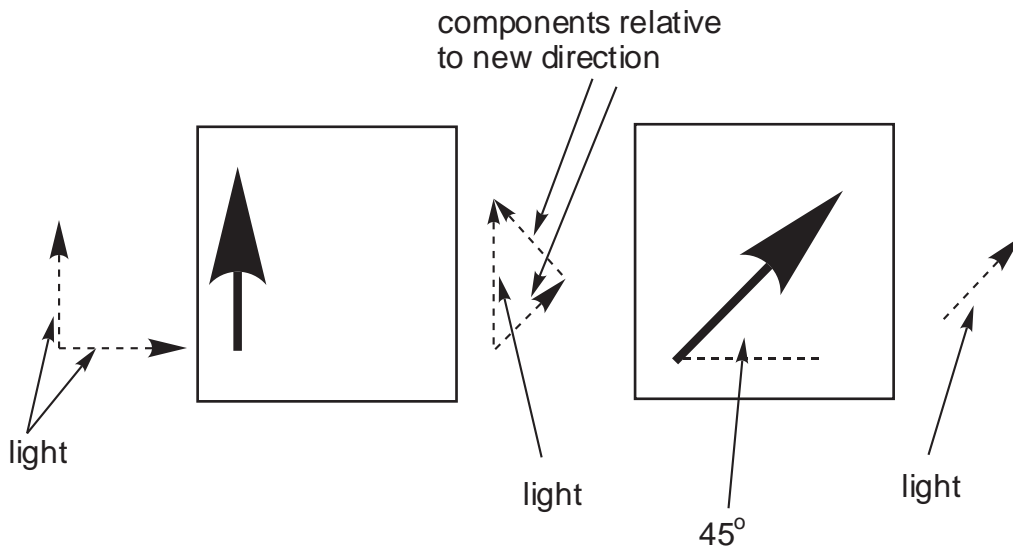
Consider 2 polaroids at right angles \rightarrow intensity out = 0. In pictures,



where the dashed arrows represent the electric field components or polarizations and the solid arrows represent the preferred directions of the polaroids. With components of vectors we have

$$\vec{E}_{in} = E_x \hat{e}_x + E_y \hat{e}_y \rightarrow \vec{E}_{intermediate} = E_y \hat{e}_y \rightarrow \vec{E}_{out} = 0$$

Now 2 polaroids at 45° degrees \rightarrow intensity out \neq 0



With components of vectors we can derive the following:

We can write the unit vector that makes an angle θ with the x-axis as

$$\hat{n} = \cos\theta\hat{e}_x + \sin\theta\hat{e}_y$$

We then have

$$\vec{E}_{in} = E_x\hat{e}_x + E_y\hat{e}_y$$

$$\begin{aligned}\vec{E}_{intermediate} &= (\vec{E}_{in} \cdot \hat{e}_y)\hat{e}_y \rightarrow \text{that is what a polaroid does} \\ &= E_y\hat{e}_y\end{aligned}$$

$$\begin{aligned}\vec{E}_{out} &= (\vec{E}_{intermediate} \cdot \hat{n})\hat{n} = ((E_y\hat{e}_y) \cdot (\cos\theta\hat{e}_x + \sin\theta\hat{e}_y))(\cos\theta\hat{e}_x + \sin\theta\hat{e}_y) \\ &= E_y \sin\theta(\cos\theta\hat{e}_x + \sin\theta\hat{e}_y)\end{aligned}$$

Some special cases are:

$$\theta = 0^\circ \rightarrow \text{orthogonal polaroids}$$

$$\vec{E}_{out} = 0$$

$$\theta = 90^\circ \rightarrow \text{parallel polaroids}$$

$$\vec{E}_{out} = E_y\hat{e}_y$$

$$\theta = 45^\circ \rightarrow \text{above example}$$

$$\vec{E}_{out} = E_y \sin 45^\circ (\cos 45^\circ \hat{e}_x + \sin 45^\circ \hat{e}_y) = \frac{E_y}{2} (\hat{e}_x + \hat{e}_y)$$

So the classical wave theory of electromagnetism seems to be able to explain polarization experiments for intense beams by using standard vector algebra.

When there are large numbers of photons in an intense beam the beam behaves as if it had wave properties. Think of water molecules making up a water wave.

Light, however, is really composed of individual particles called photons.

(4) Now let us add another polaroid to the experiment:

laser + calcite + [2 polaroids at 90 degrees] → all gone

[2 polaroids alone at 90 degrees] → all gone

This means that first polaroid reduces beams to only one direction which cannot get through the second polaroid (they are the wrong kind... have the wrong direction).

Now that things seem to be making some sort of sense, we disturb this system with another measurement.

- **if we leave the system alone NO light is observed**

- if third polaroid same as first, then same result(no change)
 there should not be any change because we are not gaining any new information....
 we already know that half is polarized in each direction
- if third polaroid same as second, then same result(no change)
 again there should not be any change since we are still not gaining any new information.....
 we already know that the remaining half are polarized in the other direction
- if third polaroid oriented in a different direction
 → different result ...
 now some light gets through the 2nd polaroid
 what does this mean?

Does it mean that we have somehow **recreated** the other kind of light?

Remove the third polaroid and put calcite in its place to see!!

It will turn out that it was the act of measurement(inserting a third polaroid to obtain new information) that disturbed the original system and changed the experimental result.

All of the physics contained in the quantum world is in these simple experiments with lasers, polaroids and calcite crystals.....we just have to pull it out and we will in our later discussions.

Let me illustrate a dilemma.

Let us start with a large number of photons (10^{23}). This very intense beam then behaves like a wave. Classical physics should then be able to explain what happens.

**Place a polaroid in a laser beam.
 Half of the light gets through.**

**Place a second polaroid at right angles.
 No light gets through.**

**Place a third polaroid in between at 45 degrees.
 Half of the light gets through.**

This is easy to explain with waves or electric field vectors and vector components as we saw earlier.

**Polaroid 1 in y-direction ($\theta = 90^\circ$ direction)
 Polaroid 2 in 45 degree direction (new y' direction or $\theta = 45^\circ$)
 Polaroid 3 in x-direction ($\theta = 0^\circ$ direction)**

We have the experiment shown below:

$$\begin{array}{ccccccc} \vec{E}_0 & | & \vec{E}_1 & | & \vec{E}_2 & | & \vec{E}_3 \\ & & \uparrow & & \uparrow & & \uparrow \\ & & \#1 & & \#2 & & \#3 & & \text{polaroids} \end{array}$$

Analyzing with vectors we have:

$$\vec{E}_0 = E_x \hat{e}_x + E_y \hat{e}_y$$

$$E_x = E_y \quad (\text{unpolarized light})$$

$$E^2 = E_x^2 + E_y^2 \rightarrow E_x = E_y = \frac{E}{\sqrt{2}}$$

$$\vec{E}_0 = \frac{E}{\sqrt{2}}(\hat{e}_x + \hat{e}_y)$$

$$\text{Energy} = E^2 \rightarrow \text{total energy in beam}$$

$$E_x^2 = \frac{E^2}{2} \rightarrow \frac{1}{2} \text{ of energy in x - polarized light}$$

$$E_y^2 = \frac{E^2}{2} \rightarrow \frac{1}{2} \text{ of energy in y - polarized light}$$

After 1

$$\vec{E}_1 = E_y \hat{e}_y = \frac{E}{\sqrt{2}} \hat{e}_y \rightarrow \text{energy} = \frac{E^2}{2} = \frac{1}{2} \text{ of original energy}$$

After 2

$$\begin{aligned} \vec{E}_2 &= \frac{E}{\sqrt{2}} \sin 45^\circ (\cos 45^\circ \hat{e}_x + \sin 45^\circ \hat{e}_y) = \frac{E}{2\sqrt{2}} (\hat{e}_x + \hat{e}_y) \\ &\rightarrow \text{energy} = \frac{E^2}{4} = \frac{1}{4} \text{ of original energy} \end{aligned}$$

After 3

$$\vec{E}_3 = \frac{E}{2\sqrt{2}} \hat{e}_x \rightarrow \text{energy} = \frac{E^2}{8} = \frac{1}{8} \text{ of original energy} \quad \text{or some light}$$

Now remove the 45 degree polaroid....

Polaroid 1 in y-direction ($\theta = 90^\circ$ direction)

Polaroid 3 in x-direction ($\theta = 0^\circ$ direction)

After 1

$$\vec{E}_1 = E_y \hat{e}_y = \frac{E}{\sqrt{2}} \hat{e}_y \rightarrow \frac{1}{2} \text{ of original energy}$$

After 3

$$\vec{E}_3 = 0 \rightarrow \text{energy} = 0 \quad \text{or no light}$$

So classical physics has no trouble explaining what is happening for intense beams where we can talk about **dividing energy** between different processes. At each step, the explanation will say that some fraction of the photons or some fraction of the energy does not pass through a particular polaroid and at each stage the intensity of the beam, which is related to the number of photons or the energy, will

change in the expected manner.

But what about any **particular** photon in the beam, each of which is polarized at 45° to the preferred axis at polaroid #3?

Now the answer is not clear at all and the fundamental dilemma of the subatomic world rears its ugly head.

Remember, each individual photon cannot be subdivided.

As will become clear during our discussions of quantum mechanics, this question about what will happen to a particular photon under certain conditions is **not very precise** and if we do not ask precise questions, then we should not be surprised that we get confusing answers or answers that seemingly defy reason.

In order for the theory to make clear predictions about experiments, we will have to learn how to ask precise questions. It will take time but we will learn how.

Remember, only questions about the results of experiments have a real significance in physics and it is only such questions that theoretical physics has to consider. The questions and the subsequent experiments devised to answer the questions must be clear and precise, however.

In this case, we can make the question clear by doing the experiment with a beam containing only one photon (very weak beam) and observe what happens as it arrives at the polaroid.

It was not until the 1980's that experiments of this sort could actually be carried out.

In particular, we make a **simple observation** to see whether or not it passes through the polaroid.

The most important result is that this single photon either passes through the polaroid or it does not.

We never observe $1/2$ the energy of a single photon passing through the polaroid. We always observe either no energy or all the energy.

One never observes a part of a photon passing through and a part getting absorbed in the polaroid.

In addition, if a photon gets through, then observation shows that its polarization **changes** such that it is now polarized in a direction parallel to the optic axis of this particular polaroid (instead of at 45° with respect to that axis as it was before it encountered the polaroid).

In a beam of N photons, each will independently behave as the single photon did. No experiment can determine which photon will pass through and which will not, even though they are all identical. In each experiment, however, exactly $1/2$ of the total energy and $1/2$ of the photons will pass through polaroid #3.

The only way this result can be interpreted is to say that each

photon has a **probability** = 1/2 of passing through. We are forced into a **probabilistic point of view** by the fact that the energy of the photons in an electromagnetic wave is **quantized (not continuous)**.

We have managed to preserve the indivisibility of the photons (or ultimately the quantization of their energy). We were able to do this **only** by abandoning the **determinacy** of classical physics with respect to identical objects, i.e., in classical physics if two objects are prepared identically, then they will behave the same way in identical experiments.

The results in this experiment are not determined by the experimental conditions (initial) under control of the experimenter, as they would have been according to classical ideas.

The most that we will be able to predict in any experiment is a set of possible results, with a probability of occurrence for each.

The experiment described above involving a single photon polarized at an angle to the optic axis, represents the only experimental and theoretical question we can ask.

It is what I shall call a **go-nogo** experiment.

Does the photon go through or is it absorbed? That is the only legitimate question we can ask in this case.

So if I arrange my experiment so that only one photon is inside the apparatus at any one time we have a **problem**. Let us redo the experiment with two polaroids at 45°.

What happens as it comes to the first polaroid?

It **either** gets through **or** it does not or **GO-NOGO**.

At the second polaroid it is again **GO-NOGO**.

In fact, **at the end**, it either gets through or it does not!

Now send another single photon through. It is identical to the first and is also **GO-NOGO**.

What they do is in no way related to each other.

Wait a century before sending the second photon so the first does its thing long before the second is even created.

What happens after a large number of unpolarized photons are sent through a single polaroid? The answer is that 1/2 get through!

Just look at the two polaroids in the experiment. If a photon gets through the first polaroid, what happens at the second?

GO-NOGO.

But if we send many, then we get exactly 1/2 of those getting through the first getting through the second (at 45°).

The only interpretation that works here is that after a photon passes through the first, the photon has a probability of 1/2 of passing through the second.

We do not know and cannot predict what **any** particular photon will do, but we can predict what a large number will do.

That is how probability works.

As we will see QM will force us to say that between the first and second polaroids the photon is in an unresolved indeterminate state with a 50-50 chance of passing through the second 45° polaroid.

This is once again the mysterious superposition rearing its ugly head!

We shall see that questions like....

What decides whether a photon goes through?
When does the photon decide whether it will pass through?
How does a photon change its polarization direction?

cannot be answered by experiment and therefore they must be regarded as outside the domain of quantum theory and possibly all of physics and cannot be relevant to any new theory we might develop.

What will our theory of quantum mechanics say about the state of the single photon?

It will be shown that the photon polarized at an angle to the optic axis is in a very **special kind** of state that we will call a **superposition** of being polarized perpendicular to the optic axis and of being polarized parallel to the optic axis.

In this state, there will exist an **extraordinary** kind of relationship between the two kinds (mutually perpendicular directions) of polarization.

The meaning of the word **superposition** follows from the **mathematical formalism and language** we are developing. It represents a new physical connection to mathematics.

This is suggested by an attempt to express the meaning of superposition in ordinary language (words). If we attempt to explain the behavior of the photon polarized at an angle to the optic axis using ordinary language, then we would have to say something like this

not polarized **parallel** to the optic axis
not polarized **perpendicular** to the optic axis
not simultaneously possessing **both** polarizations
not possessing **neither** polarization

For this experiment with only two possible polarizations, this exhausts all the logical possibilities allowed by ordinary words.

Superposition is something **completely different** than any of the above and it is **not all** of the above. Its physical content will, however, be **precise and clear** in our new mathematical formalism.

When the photon encounters the polaroid, we are observing it. We are observing whether it is polarized perpendicular or parallel to the optic axis of the polaroid. The effect of this measurement will be to end up with the photon having one or the other polarizations. It always makes a "jump" from a state of superposition to a state of a **definite** polarization.

Which of the two states it "jumps" to cannot be predicted. We can, however, predict the probability of each for a large set of identical measurements.

If it "jumps" into the parallel state, it passes through. If it "jumps" into the perpendicular state, it gets absorbed.

We will have a great deal more to say about the two new words, **superposition** and **jump**, as we proceed.

Using Photons and Polarization to Explain How Quantum Mechanics Works?

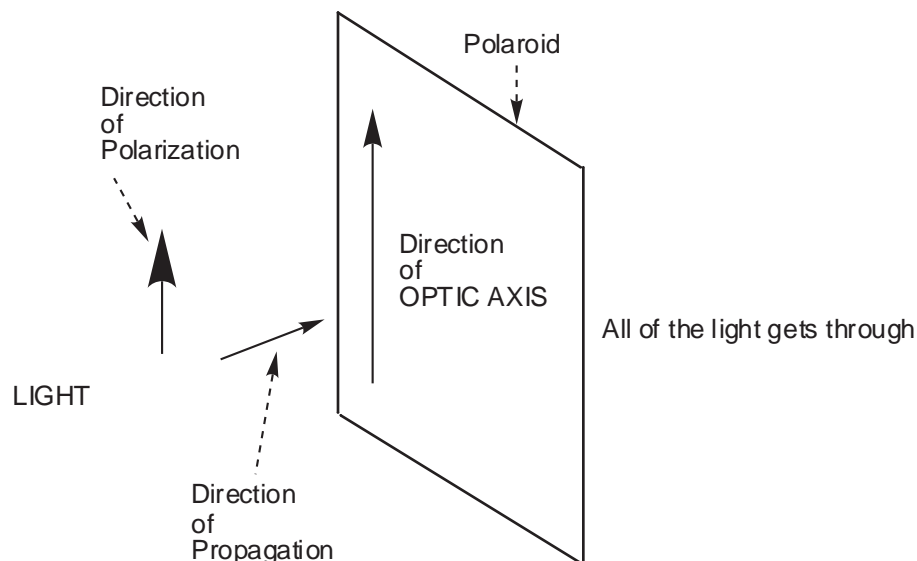
We now look at photons and polarization in more detail (repeating much of what we just said) and use our mathematical language to understand how quantum mechanics works.

As we said, many experiments indicate that electromagnetic waves have a **vector** property called **polarization**.

Suppose that we have an electromagnetic wave (we will just say **light** from now on) passing through a piece of **polaroid** material.

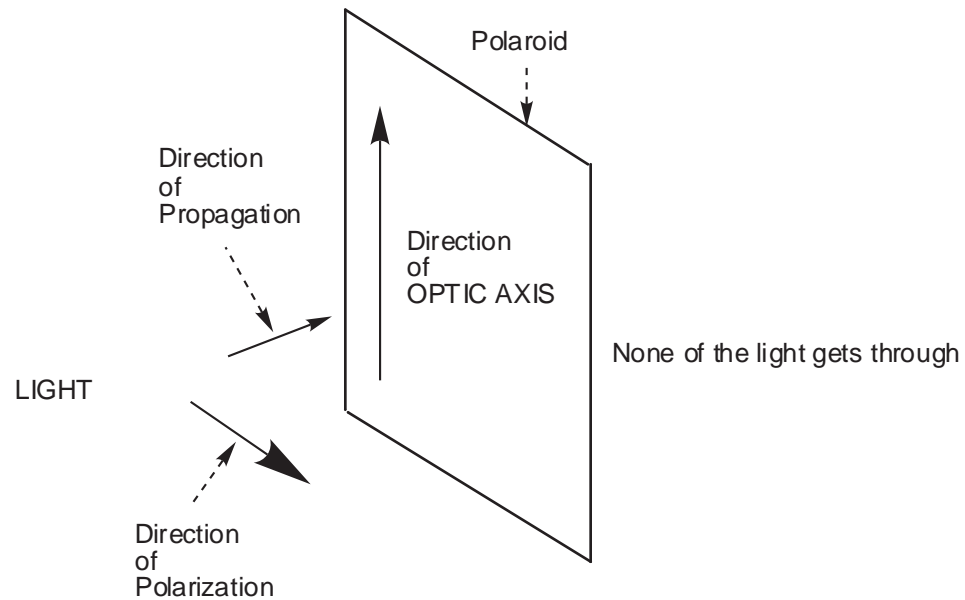
Remembering our earlier discussion, the polaroid material has the property that it only allows the light with polarization vector oriented **parallel** to a **preferred** direction in the polaroid (called the **optic axis**) to pass through the material.

Thinking classically once again, if an incident beam of light is polarized **parallel** to the optic axis (as in the figure below), then experiment says that **all** of its energy gets through the polaroid.

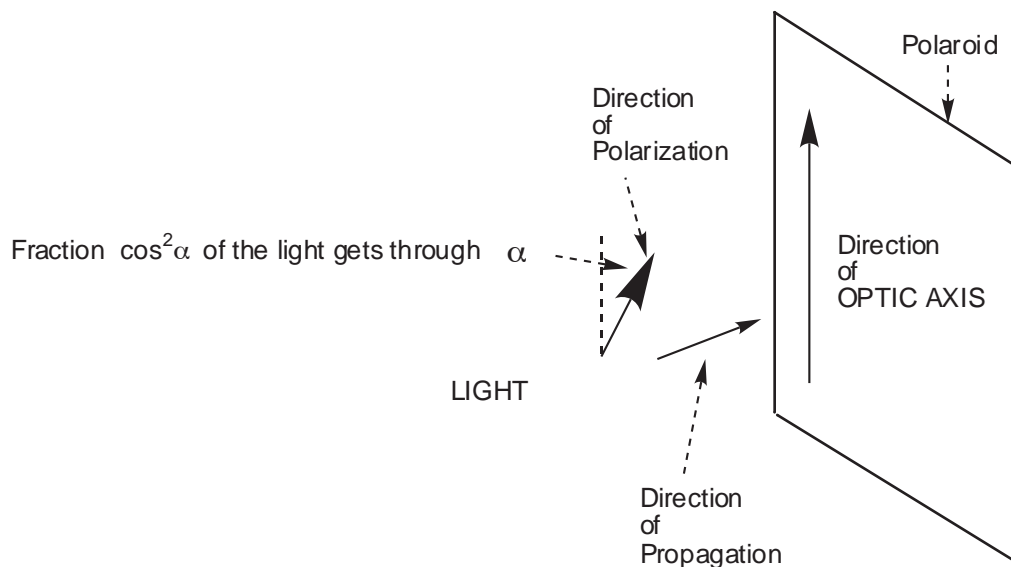


If, instead, the light is polarized **perpendicular** to the optic axis

(as in the figure below), then experiment says that **none** of its energy gets through the polaroid.



In a more general case, if it is polarized at an **angle** α to the optic axis (as in the figure below), then experiment says that a **fraction** $\cos^2\alpha$ of its energy gets through the polaroid.



By definition, when we specify the polarization of light, we are actually giving the direction of the electric field vector \vec{E} associated with the light. The polarization property or polarization vector of light depends only on the direction of the \vec{E} vector.

Classically, in Maxwell's theory light waves are represented by **plane electromagnetic waves**. This means that the associated electric field vector \vec{E} and the associated magnetic field vector \vec{B} are both perpendicular to the direction of propagation specified by a third vector \vec{k} . According to Maxwell theory, if we choose (arbitrary) the direction of propagation to be the z-axis, which is specified by the unit vector \hat{e}_z , then \vec{E} and \vec{B} lie somewhere in the x-y plane, which is the plane perpendicular to the direction of propagation. \vec{E} and \vec{B}

are perpendicular to each other

Now, any vector in the x-y plane can be specified in terms of a pair of orthonormal vectors (called the basis) in that plane. For light, the pair of orthonormal vectors are called the **basis polarization vectors**.

Two standard sets of orthonormal vectors are often chosen when one discusses polarization. One of the two sets is

$$\hat{\epsilon}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{\epsilon}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (01)$$

As we shall see, they correspond to **plane-polarized** waves.

A second orthonormal set is

$$\hat{\epsilon}_R = \hat{\epsilon}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \hat{\epsilon}_L = \hat{\epsilon}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} \quad (02)$$

They will correspond to **circularly-polarized waves**.

For classical electromagnetic fields, a light wave propagating in the z-direction is usually described (using one the two orthonormal sets) by electric field vectors of the form given below.

Plane-polarized basis:

$$\vec{E}(\vec{r}, t) = \begin{pmatrix} E_x(\vec{r}, t) \\ E_y(\vec{r}, t) \\ 0 \end{pmatrix} = E_x(\vec{r}, t) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + E_y(\vec{r}, t) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = E_x(\vec{r}, t) \hat{\epsilon}_x + E_y(\vec{r}, t) \hat{\epsilon}_y \quad (03)$$

Circular-polarized basis:

$$\vec{E}(\vec{r}, t) = \begin{pmatrix} E_x(\vec{r}, t) \\ E_y(\vec{r}, t) \\ 0 \end{pmatrix} = \left(\frac{E_x(\vec{r}, t) + iE_y(\vec{r}, t)}{\sqrt{2}} \right) \hat{\epsilon}_R + \left(\frac{E_x(\vec{r}, t) - iE_y(\vec{r}, t)}{\sqrt{2}} \right) \hat{\epsilon}_L \quad (04)$$

By convention and for mathematical simplicity, we represent the field components by

$$\begin{aligned} E_x(\vec{r}, t) &= E_x^0 e^{i(kz - \omega t + \alpha_x)} \\ E_y(\vec{r}, t) &= E_y^0 e^{i(kz - \omega t + \alpha_y)} \end{aligned} \quad (05)$$

where $k = \frac{2\pi}{\lambda}$, λ is the wavelength, ω is the angular frequency,

α_x and α_y are phases and E_x^0 and E_y^0 are the (real) amplitudes of the electric field components. This is just the traveling wave formalism we used earlier.

The actual "real" electric field components are given by the real

parts of the expressions in (05)

$$\begin{aligned}
 E_x(\vec{r}, t) &= E_x^0 e^{i(kz - \omega t + \alpha_x)} = E_x^0 \left(\cos(kz - \omega t + \alpha_x) + i \sin(kz - \omega t + \alpha_x) \right) \\
 E_{x, \text{physical}}(\vec{r}, t) &= E_x^0 \cos(kz - \omega t + \alpha_x) \\
 E_y(\vec{r}, t) &= E_y^0 e^{i(kz - \omega t + \alpha_y)} = E_y^0 \left(\cos(kz - \omega t + \alpha_y) + i \sin(kz - \omega t + \alpha_y) \right) \\
 E_{y, \text{physical}}(\vec{r}, t) &= E_y^0 \cos(kz - \omega t + \alpha_y)
 \end{aligned}
 \tag{06}$$

What do these expressions say about the physical electric field vector?

These relations say that in ideally monochromatic light (single wavelength or frequency) the x- and y-components of the electric field vector oscillate with a definite frequency at each point along the direction of propagation.

For simplicity, let us look at $z=0$. We have

$$\begin{aligned}
 E_x(\vec{r}, t) &= E_x^0 \cos(\omega t + \alpha_x) \\
 E_y(\vec{r}, t) &= E_y^0 \cos(\omega t + \alpha_y)
 \end{aligned}
 \tag{07}$$

where

$$\begin{aligned}
 \omega &= 2\pi f \\
 f &= \text{frequency} \\
 \lambda &= \text{wavelength} \\
 c &= \text{speed} = \lambda f
 \end{aligned}$$

The total electric field vector is a sum or superposition of the two components

$$\vec{E}(\vec{r}, t) = E_x(\vec{r}, t) \hat{e}_x + E_y(\vec{r}, t) \hat{e}_y
 \tag{08}$$

or it is the resultant effect produced by superposing two independent orthogonal oscillations.

Case #1: The orthogonal oscillations are **in phase**, say $\alpha_x = \alpha_y = 0$ for simplicity. We then have

$$\vec{E}(\vec{r}, t) = (E_x^0 \hat{e}_x + E_y^0 \hat{e}_y) \cos(\omega t) = \vec{E}^0 \cos(\omega t)
 \tag{09}$$

This says that the electric field vector (the tip) oscillates with the **same** frequency in a **single** direction.

This is called **linearly-polarized** or **plane-polarized light**.

Case #2: When the orthogonal oscillations are **not in phase** the resultant electric vector moves around in an ellipse, i.e., the direction is changing with time.

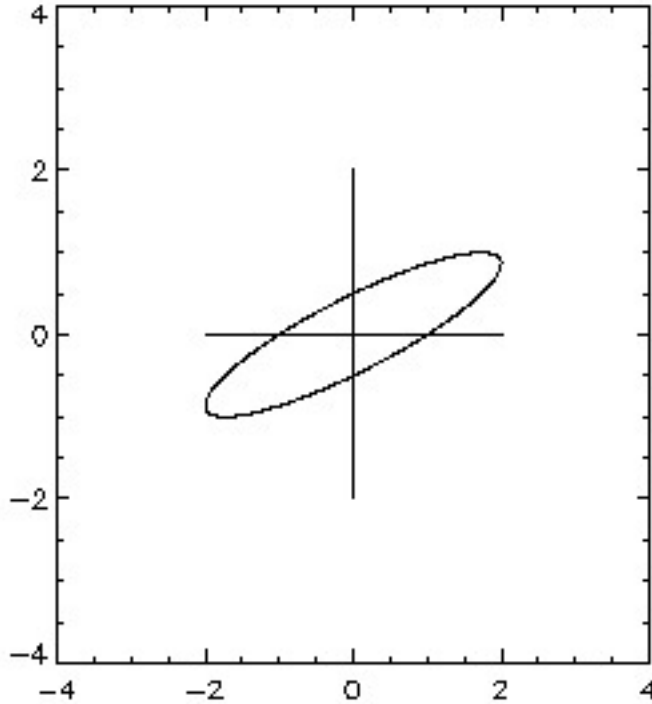
$$\begin{aligned}
 \vec{E}(\vec{r}, t) &= E_x(\vec{r}, t) \hat{e}_x + E_y(\vec{r}, t) \hat{e}_y \\
 &= E_x^0 \cos(\omega t) \hat{e}_x + E_y^0 \cos(\omega t + \alpha) \hat{e}_y
 \end{aligned}
 \tag{10}$$

where we have chosen $\alpha_x = 0, \alpha_y = \alpha$.

This corresponds to an equation for the vector tip given by

$$\left(\frac{x}{E_x^0}\right)^2 + \left(\frac{y}{E_y^0}\right)^2 - 2\cos\alpha\frac{x}{E_x^0}\frac{y}{E_y^0} = \sin^2\alpha \quad (11)$$

which looks like



for $E_x^0 = 2.0, E_y^0 = 1.0, \alpha = \pi/6$.

This is called **elliptically-polarized light**. If the ellipse is a circle, then it is called **circularly-polarized light**.

If the tip of the electric field vector, when we look at it as the light comes straight toward us, goes around in a counterclockwise direction, the light is right-hand circularly polarized. If it goes clockwise, the light is left-hand circularly polarized.

Mathematically, the relationship between the polarization state of the light and the \vec{E} vector is shown by a few examples below. Using equations (03) and (04), we have these special cases:

(1) If $E_y = 0$, the wave is **plane-polarized in the x-direction**

$$\vec{E} = E_x \hat{e}_x = E_x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (12)$$

(2) If $E_x = 0$, the wave is **plane-polarized in the y-direction**

$$\vec{E} = E_y \hat{\epsilon}_y = E_y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (13)$$

(3) If $E_x = E_y$, the wave is **plane-polarized at 45°**

$$\vec{E} = E_x \hat{\epsilon}_x + E_x \hat{\epsilon}_y = E_x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad (14)$$

(4) If $E_y = -iE_x = e^{-i\frac{\pi}{2}} E_x$, then the y-component lags the x-component by 90° (out of phase by $-\pi/2$) and the wave is **right circularly polarized**

$$\vec{E} = E_x \hat{\epsilon}_R = E_x \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \quad (15)$$

(5) If $E_y = iE_x = e^{i\frac{\pi}{2}} E_x$, then the y-component leads the x-component by 90° (out of phase by $\pi/2$) and the wave is **left circularly polarized**

$$\vec{E} = E_x \hat{\epsilon}_L = E_x \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} \quad (16)$$

This set of polarization properties is sufficient for our discussions.

One final case to consider is called **unpolarized light**.

If the relative x- and y-phase, i.e., $\alpha_x - \alpha_y = \text{relative phase}$ is not kept fixed, then the electric field vector will oscillate in different directions so that the polarization is constantly changing. In particular, if the polarization direction changes more rapidly than we can detect it, then the light is called unpolarized

Let us now return to the question - what is a polaroid?

We first discuss **birefringence**.

An interesting effect of polarization is that there are substances for which the index of refraction is different for light linearly polarized in different directions relative to the material.

Suppose some material consists of long, nonspherical molecules (much longer than they are wide) and suppose that these molecules are arranged in the material with their long axes parallel.

What happens when an oscillating electric field passes through this

material?

Suppose that because of the structure of the material, the electrons in the material respond more easily to oscillations in the direction parallel to the long axes of the molecules than they would respond if the electric field tries to push them perpendicular to the long axes. In this way we expect a different response for polarization in different directions.

The direction parallel to the long axes is called the optic axis. When the polarization is parallel to the optic axis or along the long axes of the molecules, the index of refraction is different than when the polarization is perpendicular to the optic axis. Such a material is called **birefringent**. It has two indices of refraction depending on the direction of the polarization (relative to the optic axis) inside the substance.

The index of refraction affects the electric field as it propagates through the material as shown below:

$$\begin{aligned}\vec{E}(\vec{r},t) &= \vec{E}_0 \cos(kz - \omega t) = \vec{E}_0 \cos(kz - \omega z/v) \\ &= \vec{E}_0 \cos(k - n\omega/c)z\end{aligned}\tag{17}$$

where n = index of refraction and $v = c/n$ = speed of light in material.

What do we expect to happen if we shine polarized light through a plate of birefringent material?

If the polarization is parallel to the optic axis, the light is transmitted with one velocity; if the polarization is perpendicular to the optic axis, the light is transmitted with a different velocity.

What happens when the light is linearly polarized 45° to the optic axis?

Light that is linearly polarized at 45° to the optic axis is represented by the electric field

$$\vec{E} = E_x \hat{e}_x + E_x \hat{e}_y = E_x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\tag{14}$$

or

$$\begin{aligned}\vec{E}(\vec{r},t) &= E_x^0 \cos((k - n_x \omega/c)z) \hat{e}_x + E_y^0 \cos((k - n_y \omega/c)z) \hat{e}_y \\ &= E^0 \cos 45^\circ \cos((k - n_x \omega/c)z) \hat{e}_x + E^0 \sin 45^\circ \cos((k - n_y \omega/c)z) \hat{e}_y \\ &= \frac{E^0}{\sqrt{2}} (\cos((k - n_x \omega/c)z) \hat{e}_x + \cos((k - n_y \omega/c)z) \hat{e}_y)\end{aligned}\tag{18}$$

This corresponds to representing 45° polarization as a superposition of x- and y-polarizations of equal amplitudes, frequency and in phase. Let us assume that we choose the y-axis to line up with the optic axis (x-axis is perpendicular to the optic axis). This choice is always arbitrary. As the light passes through, the phases change

at different rates, i.e., at z' the phases will not be equal

$$\left((k - n_x \omega / c)z'\right) \neq \left((k - n_y \omega / c)z'\right) \quad (19)$$

Thus, if the two components start out in phase, they will go in and out of phase as they travel through the material. The phase difference is proportional to the depth into the material, i.e.,

$$\Delta(\text{phase}) = \frac{\omega}{c} z' (n_y - n_x) \quad (20)$$

If the thickness is just right to introduce a 90° phase shift between the x- and y-components, the linearly polarized (entering material) light will leave the material circularly polarized. A plate with such a thickness is called a **quarter-wave plate**. In this case we have

$$\text{Entering} : \frac{E^0}{\sqrt{2}} (\hat{\epsilon}_x + \hat{\epsilon}_y) \cos(\omega t) \quad (21)$$

$$\text{Leaving} : \frac{E^0}{\sqrt{2}} (\cos(\omega t) \hat{\epsilon}_x + \cos(\omega t + \pi/2) \hat{\epsilon}_y) = \frac{E^0}{\sqrt{2}} (\cos(\omega t) \hat{\epsilon}_x + \sin(\omega t) \hat{\epsilon}_y)$$

If we send the light through two such quarter-wave plates, then it exits linearly polarized again in a direction at right angles to the original direction. In this case we have

$$\text{Entering} : \frac{E^0}{\sqrt{2}} (\hat{\epsilon}_x + \hat{\epsilon}_y) \cos(\omega t) \quad (22)$$

$$\text{Leaving} : \frac{E^0}{\sqrt{2}} (\cos(\omega t) \hat{\epsilon}_x + \cos(\omega t + \pi) \hat{\epsilon}_y) = \frac{E^0}{\sqrt{2}} (\hat{\epsilon}_x - \hat{\epsilon}_y) \cos(\omega t)$$

Example: Birefringence of cellophane

Cellophane consists of long, fibrous molecules. It is not isotropic since the fibers lie preferentially in one direction. We create a beam of linearly polarized light by sending unpolarized light through a sheet of polaroid. As we have said earlier, polaroid has the useful property that it transmits light that is linearly polarized parallel to preferred axis of the polaroid with very little absorption, but light polarized perpendicular to the preferred direction is strongly absorbed (not transmitted).

When the unpolarized beam is sent through the polaroid, only that part of the beam (the component of the electric field) that is vibrating parallel to the preferred axis of the polaroid gets through. In this manner, the exiting beam is linearly polarized (in the direction of the preferred axis).

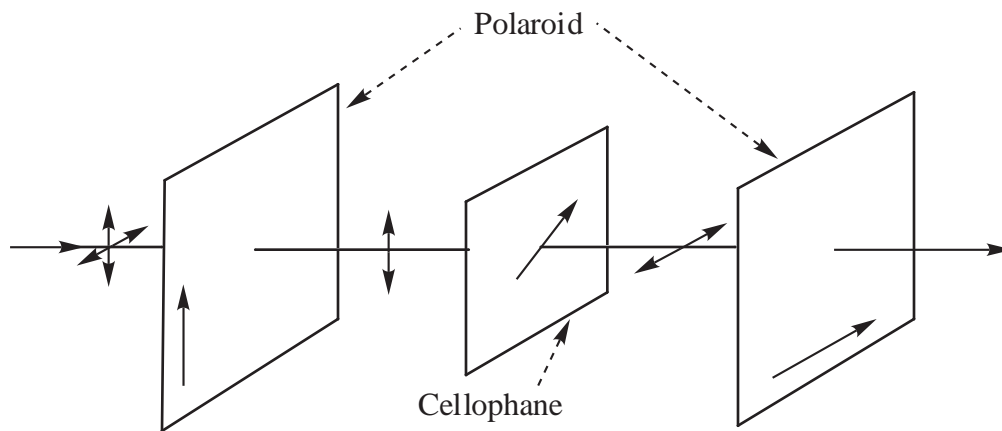
This same property of a polaroid is useful in determining the direction of polarization of any linearly polarized beam or in determining whether the beam is linearly polarized or not.

If the beam is linearly polarized, it will not be transmitted through the sheet when the preferred axis of the polaroid is orthogonal to

the direction of the polarization.

If the transmitted beam intensity is independent of the orientation of the polaroid, then the beam is not linearly polarized.

The birefringence of cellophane is illustrated by setup below:



The first polaroid produces a linearly polarized beam (from an entering unpolarized beam). The linearly polarized beam then passes through the cellophane and finally through a second polaroid. The second polaroid detects the effect of the cellophane on the polarization of the beam.

If we initially set the axes of the two polaroids orthogonal to each other, no light is transmitted through the pair (no cellophane present).

We now introduce the cellophane as shown in the diagram. If we rotate the cellophane sheet around the beam axis, we find that some light transmits through the second polaroid. In addition, there are two orthogonal orientations of the cellophane which permit no light to pass through the second polaroid.

These two directions are such that the cellophane has no effect on the linear polarization of the beam so that none gets through the second polaroid. The directions are parallel and perpendicular to the optic axis of the cellophane.

We assume that the light passes through the cellophane with two different speeds in these two different directions, but it is transmitted without changing the direction of polarization. When the cellophane is turned halfway between these two directions (as in the diagram above) we observe that the light passing through the second polaroid is bright.

It turns out that ordinary cellophane is very close to half-wave thickness for most of the colors in white light. Such a sheet will turn the direction of linear polarization through 90° if the incident linearly polarized beam makes an angle of 45° with the optic axis, so that the beam emerging from the cellophane is then vibrating in the right direction to pass through the polaroid sheet.

The cellophane will be a half-wave plate for only one wavelength in the white light and the transmitted light will be that color. The

transmitted color depends on the thickness of the cellophane.

Finally we explain **polaroids**.

Polaroids are materials where not only the index of refraction but also the amount of absorption is different for light polarized in different directions. Polaroid consists of a thin layer of small crystals of herapathite (a salt of iodine and quinine), all aligned with their axes parallel (the optic axis). These crystals absorb light when the polarization is orthogonal to this direction and do not absorb light when the polarization is parallel.

As we said earlier, if we send light into a polaroid so that the light is linearly polarized at an angle θ to the optic axis. What intensity will come through?

We simply resolve the incident light (its electric field) into components parallel ($\cos\theta$) and perpendicular ($\sin\theta$) to the optic axis. The light that comes out of the polaroid is only the $\cos\theta$ part; the $\sin\theta$ part is absorbed. The amplitude that is transmitted is smaller than the amplitude that entered.

$$\begin{aligned} \text{Entering} &: E^0(\cos\theta\hat{\epsilon}_x + \sin\theta\hat{\epsilon}_y) \\ \text{Leaving} &: E^0\cos\theta\hat{\epsilon}_x \end{aligned} \tag{23}$$

The intensity or brightness or energy is:

$$\begin{aligned} \text{Entering} &: (E^0\cos\theta)^2 + (E^0\sin\theta)^2 = (E^0)^2 \\ \text{Leaving} &: (E^0\cos\theta)^2 = (E^0)^2(\cos\theta)^2 < (E^0)^2 \end{aligned} \tag{24}$$

The absorbed intensity is $(E^0)^2(\sin\theta)^2$.

The interesting experiment (it will become a paradox later), which we have already mentioned and shall use many times in our quantum discussions later, is the following. We know that no light will be transmitted through two polaroids if their optic axes are orthogonal.

Now, however, place a third polaroid with its optic axis at 45° with respect to the first polaroid between them. Observations show that some light will now be transmitted through the second polaroid.

We know that polaroids only absorb light - they do not create light.

Nevertheless, the addition of the third polaroid allows more light get transmitted.

As before, using components of the corresponding electric fields we can explain these results easily.

Only two polaroids:

$$\begin{aligned} \text{Entering 1st polaroid} &: E^0(\cos\theta\hat{\epsilon}_{x_1} + \sin\theta\hat{\epsilon}_{y_1}) \\ \text{Leaving 1st polaroid} &: E^0\cos\theta\hat{\epsilon}_{x_2} \end{aligned}$$

Entering 2nd polaroid : $E^0 \cos\theta \hat{\epsilon}_{y_2}$

Leaving 2nd polaroid : 0

So no energy (no light) gets through

Three polaroids:

Entering 1st polaroid : $E^0 (\cos\theta \hat{\epsilon}_{x_1} + \sin\theta \hat{\epsilon}_{y_1})$

Leaving 1st polaroid : $E^0 \cos\theta \hat{\epsilon}_{x_2}$

Entering 2nd polaroid : $E^0 \cos\theta (\cos 45^\circ \hat{\epsilon}_{x_2} + \sin 45^\circ \hat{\epsilon}_{y_2})$ (26)

Leaving 2nd polaroid : $E^0 \cos\theta \cos 45^\circ \hat{\epsilon}_{x_2} = \frac{1}{\sqrt{2}} E^0 \cos\theta \hat{\epsilon}_{x_2}$

Entering 3rd polaroid : $E^0 \cos\theta \cos 45^\circ (\cos 45^\circ \hat{\epsilon}_{x_3} + \sin 45^\circ \hat{\epsilon}_{y_3})$

Leaving 3rd polaroid : $E^0 \cos\theta \cos 45^\circ \cos 45^\circ \hat{\epsilon}_{x_3} = \frac{1}{2} E^0 \cos\theta \hat{\epsilon}_{x_3}$

In this case, energy = $\frac{1}{4} (E^0 \cos\theta)^2$ gets transmitted.

We have assumed that the x-axis for each polaroid is its optic axis.

All of the preceding discussion takes place at the classical level. All the phenomena discussed can be explained with classical physics concepts.

We will see shortly that all such explanations fail at the quantum level where ideas like electric fields and components of electric fields will break down.

Many experiments, including several that we will discuss later, indicate that light (and everything else in the universe as it turns out) exhibits **both particle-like and wave-like properties**. (We will define both of these terms more carefully later and decide then if this simple statement makes any sense).

The "particle" associated with light is called a **photon** (from Einstein).

We now ask a very fundamental question: Can the experimental polarization/polaroid results indicated above be reconciled with the idea of a particle-like photon?

Suppose we assume that the classical light wave polarized in a certain direction that we have been talking about is actually made up of a large number of photons **each** of which is polarized in that same direction.

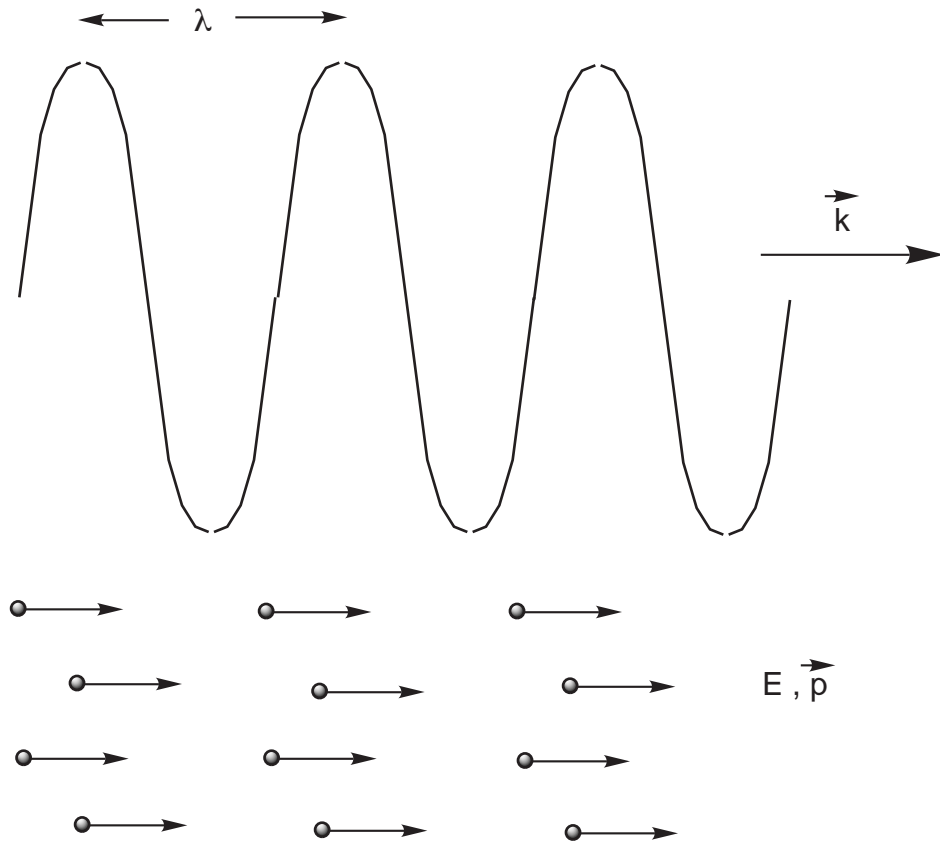
The particle properties of light, as represented by photons, invariably lead to some confusion.

It is not possible to eliminate all of this confusion at this elementary discussion level because a satisfactory treatment of

photons requires quantum electrodynamics.

We can, however, make many of the more important physical properties clear.

We consider a simple representation of an electromagnetic wave (and its associated photons) with angular frequency ω and wavelength λ moving in a direction given by the unit vector \hat{k} as shown in the figure below:



We assume that such a monochromatic electromagnetic wave is composed of N photons, each with energy E and momentum \vec{p} , such that we have the relationships

$$E = \hbar\omega \quad , \quad \vec{p} = \hbar\vec{k} = \frac{h}{\lambda} \hat{k} \quad (27)$$

where \vec{k} is the **wave vector**, $\hbar = \frac{h}{2\pi}$, $\omega = 2\pi\nu$, $h =$ Planck's constant, $\nu =$ frequency $= \frac{c}{\lambda}$, and $c =$ speed of light. We note that

$$E = \frac{h}{2\pi} 2\pi\nu = h\nu = h \frac{c}{\lambda} = pc \quad (28)$$

as required by relativity for a particle with zero mass (such as the photon).

The number of photons in the wave is such that the total energy of

the N photons, $NE = N\hbar\omega$ is equal to the total energy W in the electromagnetic wave

$$W = NE = N\hbar\omega \quad (29)$$

Here, we are using the fact, derived from many experiments, that the energy of the light wave is **quantized** and thus can only take on certain discrete values (i.e., its energy value is a multiple of some quantum of energy $\hbar\omega$). A more detailed discussion of the quantum mechanics of photon polarization will be discussed shortly.

The picture we are proposing assumes that each photon has the same polarization as the light wave, which is, in fact, verified by experiment.

This experimental property leads to some fundamental difficulties for classical mechanics.

If the incident beam is polarized parallel or perpendicular to the optic axis of a polaroid, then classical physics has no problems all the photons (and thus all the energy) either pass through or do not pass (none of the energy) through the polaroid.

But what about the case where the wave is polarized at 45° to the optic axis of the polaroid?

For the beam as a whole (a very large number of photons), the experimental result is that $1/2$ ($\cos^2 45 = 1/2$) of the total energy and hence $1/2$ of the photons pass through the polaroid.

But what about any **particular** photon, each of which is polarized at 45° to the optic axis?

Now the answer is not clear at all.

Let me repeat what I said earlier because the point is so important.

As will become clear during our discussions of quantum mechanics, this question about what will happen to a particular photon under certain conditions is not very precise.

In order for any theory to make clear predictions about experiments, we will have to learn how to ask very precise questions. We must also remember that **only questions about the results of experiments have a real significance in physics and it is only such questions that theoretical physics must consider.**

All relevant questions and the subsequent experiments devised to answer the questions must be clear and precise, however.

In this case, we can make the question clear by re-doing the experiment with a beam containing only one photon and observe what happens after it arrives at the polaroid.

In particular, we make a simple observation to see **whether or not** it passes through the polaroid.

The most important experimental result is that this single photon

either passes through the polaroid or it does not.

I will call this type of experiment a **go-no-go** experiment.

We never observe $1/2$ the energy of a single photon. We always observe either no energy or an energy exactly equal to $\hbar\omega$. One never observes a part of a photon passing through and a part getting absorbed in the polaroid.

In addition, if a photon gets through, then experiment says that its polarization vector **changes** such that it ends up polarized in a direction parallel to the optic axis of this particular polaroid (instead of at 45° with respect to that axis when it was incident on the polaroid).

In a beam of N photons, each photon will **independently** behave as the single photon did. No experiment can determine which of the photons will pass through and which will not, even though they are all identical. In each experiment, however, exactly $1/2$ of the total energy and $1/2$ of the photons will pass through the 45° polaroid.

As we shall show in this seminar, the only way this result can be interpreted is to say that each photon has a **probability** = $1/2$ of passing through the 45° polaroid.

We are forced into this probabilistic point of view due to the fact that the energy of the electromagnetic wave is quantized (or equivalently, that the electromagnetic wave is made up of photons) and we cannot have fractions of the energy quantum appearing during an experiment.

We have managed to preserve the indivisibility of the photons (the quantization of their energy), but we were able to do this only by abandoning the determinacy of classical physics and introducing probability.

The results in this experiment are not completely determined by the experimental conditions (initial) under control of the experimenter, as they would have been according to classical ideas.

As we shall see, the most that we will be able to predict in any experiment is a set of possible results, with a probability of occurrence for each.

The experiment described above involving a single photon polarized at an angle to the optic axis, allows us to ask only one type of experimental and theoretical question, namely, does the photon go through or is it absorbed?

That will turn out to be the only legitimate question we can ask in this case.

We shall see that questions like....

- What decides whether a particular photon goes through?
- When does a particular photon decide whether it will pass through?
- How does a particular photon change its polarization direction?

cannot be answered by experiment and, therefore, must be regarded as **outside the domain** of quantum theory and possibly all of physics.

Think long and hard about that last statement!!!!

What will our theory of quantum mechanics ultimately say about the state of a single photon?

It will be shown that the photon polarized at an angle to the optic axis is in a very special kind of state that we will call a **superposition** of being polarized perpendicular to the optic axis and of being polarized parallel to the optic axis, i.e., a superposition of all its possibilities (when we set up a polaroid the photons either get through or they do not - there are only two possibilities).

In this superposition state, there will exist an extraordinary kind of relationship between the two kinds (mutually perpendicular directions) of polarization. The meaning of the word **superposition** will follow from the mathematical formalism and language we will develop in this seminar. It will require a new physical connection to mathematics.

This, as we shall see later, is suggested by an attempt to express the meaning of superposition in ordinary language (words). If we attempt to explain the behavior of the photon polarized at an angle to the optic axis using ordinary language, then, as we shall see, we would have to say something like this:

The photon is

- not polarized **parallel** to the optic axis
- not polarized **perpendicular** to the optic axis
- not simultaneously possessing **both** polarizations
- not possessing **neither** polarization

For this experiment with only two possible polarizations, these statements exhaust all the logical possibilities allowed by **ordinary words** and none is correct!!!!

Superposition is something **completely different** than any of the above and it is **not all** of the above.

Its physical content will, however, be precise and clear in our new mathematical formalism.

When the photon encounters the polaroid, we are observing it. We are observing whether it is polarized perpendicular or parallel to the optic axis of the polaroid.

The effect of this measurement will be to end up with the photon having one or the other polarizations (the one we measure). In such a measurement, the photon always makes a **jump** from a state of superposition to a state of a **definite** polarization. Which of the two states it **jumps** to cannot be predicted. We can, however, predict the probability of each.

If it **jumps** into the parallel state, it has passed through. If it

jumps into the perpendicular state, it has been absorbed.

We will have a great deal more to say about the two new words, **superposition** and **jump**, as we proceed.

The Quantum Theory of Photon Polarization

We now carry out the details of a special case that will illustrate how Quantum Mechanics works and also illustrate the mathematical formalism that we developed in earlier chapters.

As we mentioned earlier, the electric field vector \vec{E} of plane electromagnetic waves lies in a plane perpendicular to the direction of propagation of the wave. Choosing the z-axis as the direction of propagation, we can represent the electric field vector as a 2-dimensional vector in the x-y plane. This means that we will only require two numbers to describe the electric field.

Since the polarization state of the light is directly related to the electric field vector, this means that we can also represent the polarization states of the photons by 2-component column vectors or **ket vectors** of the form

$$|\psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} \quad \text{where we assume the normalization condition} \quad \langle\psi|\psi\rangle = 1 \quad (30)$$

This gives the condition $|\psi_x|^2 + |\psi_y|^2 = 1$. Now we can choose to write

$$\psi_x = AE_x \quad , \quad \psi_y = AE_y \quad \text{where } A = \text{constant} \quad (31)$$

The normalization condition then implies that

$$A^2(E_x^2 + E_y^2) = 1 = A^2E^2$$

Now $E^2/8\pi$ is the energy density in an electric field and, thus, we must have

$$\text{energy of a photon} = \frac{E^2}{8\pi}V = \hbar\omega$$

where V = volume where the associated field is nonzero. This gives

$$E^2 = \frac{8\pi\hbar\omega}{V} \quad \text{or} \quad A^2 = \frac{V}{8\pi\hbar\omega}$$

or

$$\psi_x = \sqrt{\frac{V}{8\pi\hbar\omega}}E_x \quad \text{and} \quad \psi_y = \sqrt{\frac{V}{8\pi\hbar\omega}}E_y \quad (32)$$

The components in the real state vectors are independent of the volume V and depend only on the actual polarization state of the photon.

The state vector contains all the information that we can know about the state of polarization of the photon (remember that consists of just two numbers).

Examples

$$\begin{aligned}
 |x\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow x - \text{polarized photon (linear or plane polarization)} \\
 |y\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow y - \text{polarized photon (linear or plane polarization)} \\
 |R\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \rightarrow \text{Right circular - polarized photon} \\
 |L\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \rightarrow \text{Left circular - polarized photon} \\
 |45\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \text{photon polarized at } 45^\circ \text{ to the } x - \text{axis} \\
 &\quad \text{(linear or plane polarization)}
 \end{aligned} \tag{33}$$

The bra vector or linear functional corresponding the ket vector $|\psi\rangle$ is given by the row vector

$$\langle\psi| = (\psi_x^* \quad \psi_y^*) \tag{34}$$

which clearly implies via our inner product rules

$$\langle\psi|\psi\rangle = (\psi_x^* \quad \psi_y^*) \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = |\psi_x|^2 + |\psi_y|^2 = 1$$

In general, for

$$|\phi\rangle = \begin{pmatrix} \phi_x \\ \phi_y \end{pmatrix}$$

the inner product rule says

$$\langle\phi|\psi\rangle = \phi_x^* \psi_x + \phi_y^* \psi_y = \langle\psi|\phi\rangle^* \tag{35}$$

We also have

$$\langle x|x\rangle = 1 = \langle y|y\rangle \quad \text{and} \quad \langle x|y\rangle = 0 = \langle y|x\rangle \rightarrow \text{orthonormal set} \tag{36}$$

$$\langle R|R\rangle = 1 = \langle L|L\rangle \quad \text{and} \quad \langle R|L\rangle = 0 = \langle L|R\rangle \rightarrow \text{orthonormal set} \tag{37}$$

Each of these two sets is a basis for the 2-dimensional vector space of polarization states since any other state vector can be written as a linear combination of them, i.e.,

$$|\psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \psi_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \psi_x |x\rangle + \psi_y |y\rangle \tag{38}$$

$$|\psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \frac{\psi_x - i\psi_y}{2} \begin{pmatrix} 1 \\ i \end{pmatrix} + \frac{\psi_x + i\psi_y}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{\psi_x - i\psi_y}{\sqrt{2}} |R\rangle + \frac{\psi_x + i\psi_y}{\sqrt{2}} |L\rangle \quad (39)$$

We can find the components along the basis vectors using

$$\begin{aligned} \langle x|\psi\rangle &= \langle x|(\psi_x|x\rangle + \psi_y|y\rangle) = \psi_x\langle x|x\rangle + \psi_y\langle x|y\rangle = \psi_x \\ \langle y|\psi\rangle &= \langle y|(\psi_x|x\rangle + \psi_y|y\rangle) = \psi_x\langle y|x\rangle + \psi_y\langle y|y\rangle = \psi_y \end{aligned} \quad (40)$$

or

$$|\psi\rangle = |x\rangle\langle x|\psi\rangle + |y\rangle\langle y|\psi\rangle \quad (41)$$

and similarly

$$|\psi\rangle = |R\rangle\langle R|\psi\rangle + |L\rangle\langle L|\psi\rangle \quad (42)$$

Basically, we are illustrating examples of a **superposition principle**, which says that any arbitrary polarization state can be written as a superposition (linear combination) of x- and y-polarization states or equivalently, as a superposition of right- and left-circularly polarized states.

Our earlier discussions of a beam of light passing through a polaroid can now be recast in terms of these polarization states.

Classical physics says that the beam is a superposition of an x-polarized beam and a y-polarized beam and when this beam passes through an x-polaroid, its effect is to remove the y-polarized beam and pass the x-polarized beam through unchanged.

The energy of the beam is given by $|\vec{E}|^2$, which is proportional to $|\psi_x|^2 + |\psi_y|^2$. Thus, the beam energy after passing through an x-polaroid is proportional to $|\psi_x|^2$. The fraction of the beam energy or the fraction of the number of photons in the beam that passes through is given by

$$\frac{|\psi_x|^2}{|\psi_x|^2 + |\psi_y|^2} = |\psi_x|^2 = |\langle x|\psi\rangle|^2 \quad (43)$$

Our earlier discussion for the case of a single photon forced us to set this quantity equal to the probability of a single photon in the state $|\psi\rangle$ passing through an x-polaroid or

$$\begin{aligned} &\text{probability of a photon in the state } |\psi\rangle \\ &\text{passing through an x-polaroid} = |\langle x|\psi\rangle|^2 \end{aligned} \quad (44)$$

This agrees with our earlier mathematical results.

We define $\langle x|\psi\rangle$ as the **probability amplitude** for the individual photon to pass through the x-polaroid.

Another example confirming these results is light passing through a

prism.

A prism passes right-circularly-polarized(RCP) light and rejects (absorbs) left-circularly-polarized(LCP) light.

Since we can write

$$|\psi\rangle = |R\rangle\langle R|\psi\rangle + |L\rangle\langle L|\psi\rangle$$

we can generalize the polaroid result to say

$$\langle R|\psi\rangle = \text{amplitude that a photon in state } |\psi\rangle \text{ passes through the prism}$$

and

$$|\langle R|\psi\rangle|^2 = \text{probability that a photon in state } |\psi\rangle \text{ passes through the prism}$$

(45)

Polaroids and prisms are examples of **go-nogo** devices. Certain photons are passed through while others are absorbed in these devices.

How Many Basis Sets?

We have already seen two examples of basis sets for the 2-dimensional vector space of polarization states, namely,

$$\{|x\rangle, |y\rangle\} \text{ and } \{|R\rangle, |L\rangle\}$$

In the 2-dimensional vector space there are an infinite number of such orthonormal basis sets related to the $\{|x\rangle, |y\rangle\}$ set. They are all equivalent for describing physical systems (they correspond to different orientations of the polaroid in the experimental measurement).

We can obtain the other sets say $\{|x'\rangle, |y'\rangle\}$, by a rotation of the bases (or axes) as shown in Figure 1a.

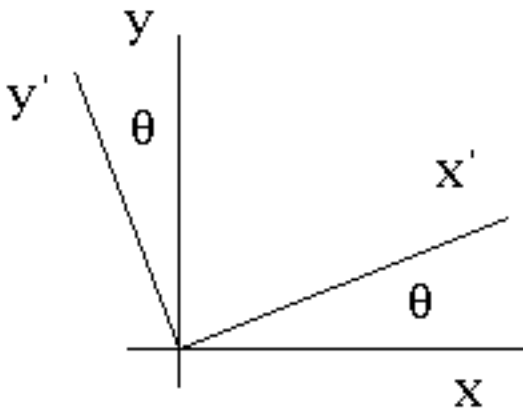


Figure 1a

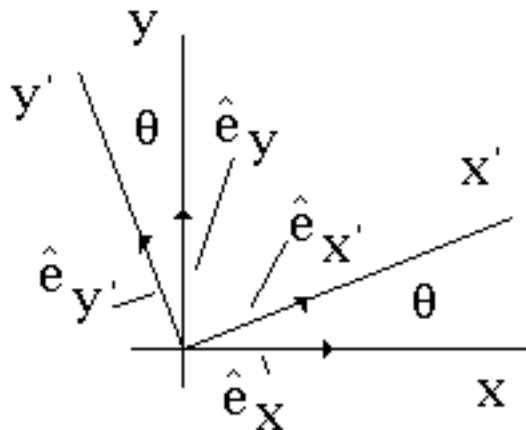


Figure 1b

We then have in the x-y basis

$$|\psi\rangle = \psi_x|x\rangle + \psi_y|y\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \begin{pmatrix} \langle x|\psi\rangle \\ \langle y|\psi\rangle \end{pmatrix} \quad (46)$$

and if we choose to use the equivalent $x'-y'$ basis we have

$$|\psi\rangle = \psi_{x'}|x'\rangle + \psi_{y'}|y'\rangle = \begin{pmatrix} \psi_{x'} \\ \psi_{y'} \end{pmatrix} = \begin{pmatrix} \langle x'|\psi\rangle \\ \langle y'|\psi\rangle \end{pmatrix} \quad (47)$$

How are these components related to each other?

We have from equation (41) that

$$|\psi\rangle = |x\rangle\langle x|\psi\rangle + |y\rangle\langle y|\psi\rangle$$

which implies

$$\begin{aligned} \langle x'|\psi\rangle &= \langle x'|x\rangle\langle x|\psi\rangle + \langle x'|y\rangle\langle y|\psi\rangle \\ \langle y'|\psi\rangle &= \langle y'|x\rangle\langle x|\psi\rangle + \langle y'|y\rangle\langle y|\psi\rangle \end{aligned} \quad (48)$$

or in matrix notation

$$\begin{pmatrix} \langle x'|\psi\rangle \\ \langle y'|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle x'|x\rangle & \langle x'|y\rangle \\ \langle y'|x\rangle & \langle y'|y\rangle \end{pmatrix} \begin{pmatrix} \langle x|\psi\rangle \\ \langle y|\psi\rangle \end{pmatrix} \quad (49)$$

So we can transform the basis (transform the components) if we can determine the 2x2 transformation matrix in equation (49).

It turns out that this result is quite general in the sense that it holds for any two bases, not just the linear polarized bases used to derive it.

For the linear(plane) polarized case, we can think of an analogy to unit vectors along the axes in ordinary space as shown in Figure 1b. Then we have(by analogy)

$$\begin{aligned} \hat{e}_x \cdot \hat{e}_{x'} &= \cos\theta = \langle x'|x\rangle \quad , \quad \hat{e}_x \cdot \hat{e}_{y'} = \sin\theta = \langle x'|y'\rangle \\ \hat{e}_x \cdot \hat{e}_{y'} &= \cos\theta = \langle y'|y\rangle \quad , \quad \hat{e}_{y'} \cdot \hat{e}_x = -\sin\theta = \langle y'|x\rangle \end{aligned} \quad (50)$$

or

$$\begin{aligned} |x\rangle &= \langle x|x'\rangle|x'\rangle + \langle x|y'\rangle|y'\rangle = \cos\theta|x'\rangle - \sin\theta|y'\rangle \\ |y\rangle &= \langle y|x'\rangle|x'\rangle + \langle y|y'\rangle|y'\rangle = \sin\theta|x'\rangle + \cos\theta|y'\rangle \end{aligned} \quad (51)$$

or

$$\begin{pmatrix} \langle x'|\psi\rangle \\ \langle y'|\psi\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \langle x|\psi\rangle \\ \langle y|\psi\rangle \end{pmatrix} \quad (52)$$

with the transformation matrix, $\hat{R}(\theta)$ given by

$$\hat{R}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (53)$$

There are two equivalent ways to interpret the result in eq (52).

First, we could say it tells us the components of $|\psi\rangle$ in the rotated basis (we keep the vector fixed and rotate the axes).

Second, we can rotate the vector and keep the axes fixed (rotate in the opposite direction). In this case, we regard

$$\begin{pmatrix} \langle x' | \psi \rangle \\ \langle y' | \psi \rangle \end{pmatrix}$$

as a new vector $|\psi'\rangle$ whose components in the fixed x-y basis are the same as the components of $|\psi\rangle$ in the x'-y'-basis or

$$\langle x' | \psi \rangle = \langle x | \psi' \rangle \quad \text{and} \quad \langle y' | \psi \rangle = \langle y | \psi' \rangle \quad (54)$$

For real ψ_x and ψ_y , $|\psi'\rangle$ is the vector $|\psi\rangle$ rotated clockwise by θ or, regarding $\hat{R}(\theta)$ as a linear operator in the vector space we have

$$|\psi'\rangle = \hat{R}(\theta)|\psi\rangle \quad (55)$$

It is a transformation of vectors and is a unitary operator. We can see this as follows:

$$\hat{R}^{-1}(\theta) = \hat{R}(-\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \hat{R}^T(\theta) = \hat{R}^\dagger(\theta) \quad (56)$$

Transformation operators are unitary because they transform state vectors which must not change lengths (otherwise probability ideas are messed up). This follows from the fact that unitary transformations preserve inner products

$$\begin{aligned} |\psi'\rangle &= \hat{R}(\theta)|\psi\rangle, \quad |\phi'\rangle = \hat{R}(\theta)|\phi\rangle \\ \langle\phi'|\psi'\rangle &= \langle\phi|\hat{R}^\dagger(\theta)\hat{R}(\theta)|\psi\rangle = \langle\phi|\hat{R}^{-1}(\theta)\hat{R}(\theta)|\psi\rangle = \langle\phi|\hat{I}|\psi\rangle = \langle\phi|\psi\rangle \end{aligned} \quad (57)$$

Since $\hat{R}(\theta)$ is a unitary transformation operator for rotations a very general theorem says that we can express it as an exponential operator involving the angular momentum operator with respect to the axis of rotation (z-axis) \hat{J}_z , of the form

$$\hat{R}(\theta) = e^{i\theta \frac{\hat{J}_z}{\hbar}} \quad (58)$$

PROOF: We can rewrite $\hat{R}(\theta)$ as

$$\begin{aligned} \hat{R}(\theta) &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \cos\theta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\sin\theta \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ &= \cos\theta \hat{I} + i\sin\theta \hat{Q} \end{aligned} \quad (59)$$

where the physical meaning of the operator \hat{Q} is yet to be determined.

We now show that $\hbar\hat{Q} = \hat{J}_z$. Expanding equation (58) in a power series we have

$$\begin{aligned}\hat{R}(\theta) &= e^{i\theta\frac{\hat{J}_z}{\hbar}} = \hat{I} + (i\frac{\hat{J}_z}{\hbar})\theta + \frac{(i\frac{\hat{J}_z}{\hbar})^2}{2!}\theta^2 + \frac{(i\frac{\hat{J}_z}{\hbar})^3}{3!}\theta^3 + \frac{(i\frac{\hat{J}_z}{\hbar})^4}{4!}\theta^4 + \dots \\ &= \hat{R}(0) + \frac{1}{1!}\frac{d\hat{R}(\theta)}{d\theta}\Big|_{\theta=0}\theta + \frac{1}{2!}\frac{d^2\hat{R}(\theta)}{d\theta^2}\Big|_{\theta=0}\theta^2 + \frac{1}{3!}\frac{d^3\hat{R}(\theta)}{d\theta^3}\Big|_{\theta=0}\theta^3 + \frac{1}{4!}\frac{d^4\hat{R}(\theta)}{d\theta^4}\Big|_{\theta=0}\theta^4 + \dots\end{aligned}\quad (60)$$

Using $\hbar\hat{Q} = \hat{J}_z$, we have $\hat{J}_z^2 = \hbar^2\hat{I}$ so that we can write

$$\begin{aligned}\hat{R}(\theta) &= e^{i\theta\frac{\hat{J}_z}{\hbar}} = \hat{I} + (i)\frac{\hat{J}_z}{\hbar}\theta + \frac{(i)^2\hat{I}}{2!}\theta^2 + \frac{(i)^3\frac{\hat{J}_z}{\hbar}}{3!}\theta^3 + \frac{(i)^4\hat{I}}{4!}\theta^4 + \dots \\ &= (1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots)\hat{I} + i(\theta - \frac{\theta^3}{3!} + \dots)\frac{\hat{J}_z}{\hbar} \\ &= \cos\theta\hat{I} + i\sin\theta\hat{Q}\end{aligned}$$

which agrees with equation (58). Thus, we have

$$\hat{J}_z = \hbar\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\quad (61)$$

must be the matrix representing the operator \hat{J}_z in the $|x\rangle, |y\rangle$ basis.

We now work out the eigenvectors and eigenvalues of $\hat{R}(\theta)$, which are given by the equation

$$\hat{R}(\theta)|\psi\rangle = c|\psi\rangle\quad (62)$$

where c = the eigenvalue corresponding to the eigenvector $|\psi\rangle$. Since all vectors are eigenvectors of the identity operator \hat{I} , we only need to find the eigenvectors and eigenvalues of \hat{J}_z in order to solve the problem for $\hat{R}(\theta)$. We let

$$\hat{J}_z|\psi\rangle = \lambda|\psi\rangle\quad (63)$$

Now, since $\hat{J}_z^2 = \hbar^2\hat{I}$ we have

$$\hat{J}_z^2|\psi\rangle = \lambda^2|\psi\rangle = \hbar^2\hat{I}|\psi\rangle = \hbar^2|\psi\rangle$$

which says that

$$\lambda^2 = \hbar^2 \text{ or } \lambda = \pm\hbar = \text{eigenvalues of } \hat{J}_z$$

We can find the corresponding eigenvectors by inserting the eigenvalues into equation (63)

$$\hat{J}_z|J_z = \hbar\rangle = \hbar|J_z = \hbar\rangle$$

We assume that

$$|J_z = \hbar\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \text{ where } |a|^2 + |b|^2 = 1$$

to get

$$\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \hbar \begin{pmatrix} a \\ b \end{pmatrix} = \hbar \begin{pmatrix} -ib \\ ia \end{pmatrix}$$

This gives the result $ia = b$, which together with the normalization condition says that $a = \frac{1}{\sqrt{2}}$. We have arbitrarily chosen a to be real since only the relative phase between components will be important in quantum mechanics. This then gives $b = \frac{i}{\sqrt{2}}$. Finally we have the eigenvector

$$|J_z = \hbar\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = |R\rangle \quad (64)$$

Similarly, we get

$$|J_z = -\hbar\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = |L\rangle \quad (65)$$

So the eigenvectors of \hat{J}_z and hence of $\hat{R}(\theta)$ are the RCP and LCP basis states. We then have

$$\begin{aligned} \hat{R}(\theta)|R\rangle &= (\cos\theta\hat{1} + i\sin\theta\frac{\hat{J}_z}{\hbar})|R\rangle \\ &= (\cos\theta + i\sin\theta)|R\rangle \\ &= e^{i\theta}|R\rangle \\ \hat{R}(\theta)|L\rangle &= e^{-i\theta}|L\rangle \end{aligned} \quad (66)$$

Physically, this says that the $|R\rangle$ and $|L\rangle$ states are only changed by an overall phase factor under rotation of the basis. This allows us to specify what happens to an arbitrary vector $|\psi\rangle$ under rotations.

First, we expand the arbitrary vector in the $|R\rangle, |L\rangle$ basis.

$$|\psi\rangle = |R\rangle\langle R|\psi\rangle + |L\rangle\langle L|\psi\rangle \quad (67)$$

We then apply the rotation operator to obtain

$$\begin{aligned} \hat{R}(\theta)|\psi\rangle &= \hat{R}(\theta)|R\rangle\langle R|\psi\rangle + \hat{R}(\theta)|L\rangle\langle L|\psi\rangle \\ &= e^{i\theta}|R\rangle\langle R|\psi\rangle + e^{-i\theta}|L\rangle\langle L|\psi\rangle \end{aligned} \quad (68)$$

or the RCP component is multiplied by the phase factor $e^{i\theta}$ and the LCP component is multiplied by a **different** phase factor $e^{-i\theta}$.

Thus, rotations change the **relative** phase of the components, which is a **real** physical change (as opposed to an overall phase change of the state vector).

Now, it is an experimental fact that if a photon traveling in the z-direction is absorbed by matter, then the z-component of the angular momentum of the absorber increases by \hbar or decreases by \hbar . It never remains the same, nor does it change by any value other than $\pm\hbar$.

We interpret these results to say that the RCP photon is in a state which is an eigenvector of \hat{J}_z with eigenvalue \hbar or that the photon in that state has spin = \hbar . Similarly, a LCP photon has spin = $-\hbar$.

One cannot predict, for any single photon, whether the change will be \hbar or $-\hbar$. We can, however, predict the probability of either value occurring. In particular, according to our probability formalism, we must have

$$\begin{aligned} |\langle R|\psi\rangle|^2 &= \text{probability of } +\hbar \\ |\langle L|\psi\rangle|^2 &= \text{probability of } -\hbar \end{aligned} \quad (69)$$

and the average value of the z-component of the angular momentum is

$$\langle \hat{J}_z \rangle = \sum_{\text{all possibilities}} (\text{eigenvalue}) \times (\text{probability of the eigenvalue})$$

or

$$\langle \hat{J}_z \rangle = \hbar |\langle R|\psi\rangle|^2 - \hbar |\langle L|\psi\rangle|^2 \quad (70)$$

In general, a photon is neither pure RCP nor pure LCP and the angular momentum does not have a definite value.

We can still talk in terms of probabilities, however.

The **discreteness** of the angular momentum spectrum forces a probabilistic interpretation on us.

We can easily see how all of this works using our mathematical formalism as follows:

$$\begin{aligned} \langle \hat{J}_z \rangle &= \langle \psi | \hat{J}_z | \psi \rangle \\ | \psi \rangle &= | R \rangle \langle R | \psi \rangle + | L \rangle \langle L | \psi \rangle \\ \langle \hat{J}_z \rangle &= (\langle R | \psi \rangle^* \langle R | + \langle L | \psi \rangle^* \langle L |) \hat{J}_z (| R \rangle \langle R | \psi \rangle + | L \rangle \langle L | \psi \rangle) \\ &= \langle R | \hat{J}_z | R \rangle |\langle R | \psi \rangle|^2 + \langle L | \hat{J}_z | L \rangle |\langle L | \psi \rangle|^2 + \\ &\quad \langle R | \hat{J}_z | L \rangle \langle R | \psi \rangle^* \langle L | \psi \rangle + \langle L | \psi \rangle^* \langle R | \psi \rangle \langle L | \hat{J}_z | R \rangle \\ &= \hbar |\langle R | \psi \rangle|^2 - \hbar |\langle L | \psi \rangle|^2 \end{aligned}$$

as in (70).

Let us return for a moment to the matrix representation of the \hat{J}_z operator. We have found the following results:

$$\hat{J}_z | R \rangle = +\hbar | R \rangle \quad \text{and} \quad \hat{J}_z | L \rangle = -\hbar | L \rangle$$

In the $\{|R\rangle, |L\rangle\}$ basis, these relations imply the matrix representation

$$\hat{J}_z = \begin{pmatrix} \langle R|\hat{J}_z|R\rangle & \langle R|\hat{J}_z|L\rangle \\ \langle L|\hat{J}_z|R\rangle & \langle L|\hat{J}_z|L\rangle \end{pmatrix} = \hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which is the standard form of \hat{J}_z in terms of one of the so-called **Pauli matrices**, namely

$$\hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \rightarrow \hat{J}_z = \hbar \hat{\sigma}_z \quad (71)$$

Now

$$|x\rangle = \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle) \quad \text{and} \quad |y\rangle = \frac{i}{\sqrt{2}}(|R\rangle - |L\rangle)$$

and, therefore, in the $\{|x\rangle, |y\rangle\}$ basis we have the matrix representation

$$\hat{J}_z = \begin{pmatrix} \langle x|\hat{J}_z|x\rangle & \langle x|\hat{J}_z|y\rangle \\ \langle y|\hat{J}_z|x\rangle & \langle y|\hat{J}_z|y\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

which is the form we guessed earlier.

Projection Operators

Let us now turn our attention to two other important operators, namely, **projection operators** and **density operators**, in the context of photon polarization.

The projection operator $|\psi\rangle\langle\phi|$ can be represented by a 2x2 matrix in the polarization state vector space. It is constructed using the mathematical object called an **outer product**.

$$\hat{P} = |\psi\rangle\langle\phi| = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} \begin{pmatrix} \phi_x^* & \phi_y^* \end{pmatrix} = \begin{pmatrix} \psi_x\phi_x^* & \psi_x\phi_y^* \\ \psi_y\phi_x^* & \psi_y\phi_y^* \end{pmatrix} \quad (72)$$

or **equivalently**, by choosing a basis and finding the matrix representation

$$\begin{aligned} \hat{P} &= \begin{pmatrix} \langle x|\hat{P}|x\rangle & \langle x|\hat{P}|y\rangle \\ \langle y|\hat{P}|x\rangle & \langle y|\hat{P}|y\rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle x|\psi\rangle\langle\phi|x\rangle & \langle x|\psi\rangle\langle\phi|y\rangle \\ \langle y|\psi\rangle\langle\phi|x\rangle & \langle y|\psi\rangle\langle\phi|y\rangle \end{pmatrix} \\ &= \begin{pmatrix} \psi_x\phi_x^* & \psi_x\phi_y^* \\ \psi_y\phi_x^* & \psi_y\phi_y^* \end{pmatrix} \end{aligned} \quad (73)$$

In particular, we have

$$|x\rangle\langle x| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad |x\rangle\langle y| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad |y\rangle\langle x| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad |y\rangle\langle y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (74)$$

From these results we easily see that

$$|x\rangle\langle x| + |y\rangle\langle y| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \hat{I}$$

and

$$|\psi\rangle = \hat{I}|\psi\rangle = |x\rangle\langle x|\psi\rangle + |y\rangle\langle y|\psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

as we specified in our earlier mathematical discussions. Similarly, we have

$$\hat{I} = |R\rangle\langle R| + |L\rangle\langle L|$$

which leads to

$$\hat{J}_z = \hat{J}_z \hat{I} = \hat{J}_z |R\rangle\langle R| + \hat{J}_z |L\rangle\langle L| = \hbar |R\rangle\langle R| - \hbar |L\rangle\langle L| \quad (75)$$

which is the expansion of the operator \hat{J}_z in terms of eigenvalues and 1-dimensional subspace projection operators (eigenvectors) that we discussed earlier.

The action of a polarizer can be considered as a measurement. What are the operators representing such measurements? Clearly, the operators for x- and y-polarizers are given by

$$\hat{O}_x = |x\rangle\langle x| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{O}_y = |y\rangle\langle y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

since

$$\hat{O}_x |\psi\rangle = \hat{O}_x (a|x\rangle + b|y\rangle) = (|x\rangle\langle x|)(a|x\rangle + b|y\rangle) = a|x\rangle$$

and so on.

If light is polarized at an angle θ from the x-axis, it is in the state

$$|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

The operator representing the polarizer at angle θ is (in the x-y basis)

$$\hat{O}_\theta = |\theta\rangle\langle\theta| = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}$$

Note that the probability of measuring x-polarization when in a θ state is

$$\langle x|\hat{O}_\theta|x\rangle = \langle x|\theta\rangle\langle\theta|x\rangle = |\langle x|\theta\rangle|^2 = \cos^2\theta$$

as expected.

Amplitudes and Probabilities

The **probability interpretation** we have been making follows from the concept of **superposition**.

The superposition idea says that we can write any arbitrary photon state as a linear combination of basis states

$$|\psi\rangle = |R\rangle\langle R|\psi\rangle + |L\rangle\langle L|\psi\rangle$$

and then interpret $|\langle R|\psi\rangle|^2$ as the probability that the photon in the state $|\psi\rangle$ will behave as a RCP photon in the state $|R\rangle$.

Generalizing this statement, we say that a system in a state $|\psi\rangle$, in Quantum Mechanics, has a probability $|\langle\phi|\psi\rangle|^2$ of behaving like it was in the state $|\phi\rangle$.

You might now conclude, from the experimental fact that only $\pm\hbar$ is transferred to matter, that photons are **always** either in the state $|R\rangle$ with probability α **or** in the state $|L\rangle$ with probability $1-\alpha$.

This cannot be correct, however, as we can see by the following arguments.

FACT: An x-polarized photon **never** passes through a y-polaroid

PROBLEM: If, the above interpretation of being either $|R\rangle$ or $|L\rangle$ was true, then

(a) an x-polarized photon has a probability $= |\langle R|x\rangle|^2 = \frac{1}{2}$ of being RCP

and a RCP photon has a probability $= |\langle y|R\rangle|^2 = \frac{1}{2}$ of being a y-polarized photon and thus passing through a y-polaroid.

(b) an x-polarized photon has a probability $= |\langle L|x\rangle|^2 = \frac{1}{2}$ of being LCP

and a LCP photon has a probability $= |\langle y|L\rangle|^2 = \frac{1}{2}$ of being a y-polarized photon and thus passing through a y-polaroid.

This means that the total probability that an x-polarized photon would get through a y-polaroid in this interpretation is

$$\text{total probability} = |\langle R|x\rangle|^2 |\langle y|R\rangle|^2 + |\langle L|x\rangle|^2 |\langle y|L\rangle|^2 = \frac{1}{2} \quad (76)$$

However, as we stated, it **NEVER HAPPENS**. What is wrong?

SOLUTION:

When we think of an x-polarized photon being a RCP photon or a LCP photon with equal probability, we are ruling out the possibility of any **interference** effects between the RCP and LCP amplitudes.

We give meaning to the word **interference** here in this way.

The correct calculation of the probability, which lays the groundwork for all of the amplitude mechanics in Quantum Mechanics, goes as follows:

- (a) The probability amplitude of an x-polarized photon passing through a y-polaroid = $\langle y|x \rangle = 0$, which implies that the probability = $|\langle y|x \rangle|^2 = 0$ also.
- (b) If we say that the x-polarized photon is in a superposition of $|R\rangle$ and $|L\rangle$ (we make **no** statement about probabilities at this point), this implies that

$$|x\rangle = |R\rangle\langle R|x\rangle + |L\rangle\langle L|x\rangle$$

which gives

$$\langle y|x\rangle = \langle y|R\rangle\langle R|x\rangle + \langle y|L\rangle\langle L|x\rangle$$

or the **amplitude** for an x-polarized photon to pass through a y-polaroid is the **sum of two amplitudes**, namely, that it passes through as a RCP photon $\langle y|R\rangle\langle R|x\rangle$ and that it passes through as a LCP photon $\langle y|L\rangle\langle L|x\rangle$

- (c) The **probability** of passing through is then the **absolute square of the total amplitude**

$$\begin{aligned} \text{probability} &= |\langle y|R\rangle\langle R|x\rangle + \langle y|L\rangle\langle L|x\rangle|^2 \\ &= (\langle y|R\rangle^* \langle R|x\rangle^* + \langle y|L\rangle^* \langle L|x\rangle^*)(\langle y|R\rangle\langle R|x\rangle + \langle y|L\rangle\langle L|x\rangle) \\ &= |\langle y|R\rangle|^2 |\langle R|x\rangle|^2 + |\langle y|L\rangle|^2 |\langle L|x\rangle|^2 \\ &\quad + \langle y|R\rangle\langle R|x\rangle\langle y|L\rangle^* \langle L|x\rangle^* + \langle y|R\rangle^* \langle R|x\rangle\langle y|L\rangle\langle L|x\rangle \end{aligned}$$

- (d) The first two terms are the same as the incorrect calculation in equation (76). The last two terms represent **interference** effects between the two amplitudes (RCP way and LCP way).

A simple calculation shows that the interference terms exactly cancel the first two terms and that the probability equals zero in agreement with experiment !!

INTERPRETATION:

The way to interpret this result is as follows:

$\langle y|R\rangle\langle R|x\rangle$ = probability amplitude for an x-polarized photon to pass through a y-polaroid as a RCP photon

$\langle y|L\rangle\langle L|x\rangle$ = probability amplitude for an x-polarized photon to pass through a y-polaroid as a LCP photon

These are **indistinguishable** ways for the process to occur, i.e., no

measurement exists that can tell us whether it passes though as an RCP photon or as a LCP photon without destroying the interference, i.e., without radically altering the experiment.

To get the correct total probability, we add all the amplitudes for indistinguishable ways and then square the resulting total amplitude.

In the incorrect calculation, we found the probability for each indistinguishable way and then added the probabilities.

In one case, we eliminated the interference effects and got the wrong result and, in the other case, we included the interference effects and obtained the correct result.

Summarizing, we have these **rules for amplitude mechanics and probabilities** in Quantum Mechanics:

- (1) The probability amplitude for two successive events is the **product** of the amplitudes for each event, i.e., the amplitude for the x-polarized photon to pass through the y-polaroid as a RCP polarized photon is the product of the amplitude for x-polarized photon to be a RCP photon $\langle R|x \rangle$ and the amplitude for a RCP photon to be a y-polarized photon $\langle y|R \rangle$

$$\langle R|x \rangle \langle y|R \rangle$$

- (2) The total amplitude for a process that can take place in several **indistinguishable** ways is the **sum** of the amplitudes for each individual way, i.e.,

$$\langle y|x \rangle = \langle y|R \rangle \langle R|x \rangle + \langle y|L \rangle \langle L|x \rangle$$

We note here that this is merely a reflection of the property of projection operators that

$$\hat{I} = |R\rangle\langle R| + |L\rangle\langle L|$$

which says that

$$\langle y|x \rangle = \langle y|\hat{I}|x \rangle = \langle y|R \rangle \langle R|x \rangle + \langle y|L \rangle \langle L|x \rangle$$

Thus, the mathematical sum over all projection operators being equal to the identity operator is physically equivalent to the sum over all possible intermediate states and it turns into a sum over all the amplitudes for indistinguishable ways in this interpretation.

- (3) The total probability for the process to occur is the absolute square of the total amplitude.

So, in classical physics, we

- (1) find amplitudes and probabilities of each way separately
- (2) add all probabilities to get total probability

We get **NO** interference effects!!

In Quantum Mechanics, we

- (1) find the amplitudes for each indistinguishable way the process can occur
- (2) add all the amplitudes to get a total amplitude
- (3) square the total amplitude to get the total probability

We get interference effects!!

The important result here is that we must consider **ALL INDISTINGUISHABLE WAYS** in step (2).

An indistinguishable way is characterized as follows:

- (1) If two ways are indistinguishable, then there exists **no measurement** that can decide which of the two ways actually happened without altering the experiment.
- (2) In particular, if we attempt to find out, then the interference effects will disappear and we will return to the classical result obtained by adding probabilities.

What actually happens is that during any measurement trying to distinguish the ways, the relative phase of the components in the superposition becomes completely uncertain and this will wash out the interference.

This happens as follows: instead of

$$|x\rangle = |R\rangle\langle R|x\rangle + |L\rangle\langle L|x\rangle$$

if we attempted to add a measurement to determine of the x-polarized photon was RCP or LCP, we would have

$$|\tilde{x}\rangle = e^{i\alpha_R}|R\rangle\langle R|x\rangle + e^{i\alpha_L}|L\rangle\langle L|x\rangle$$

The probability calculation then gives

$$\begin{aligned} \text{total probability} = & |\langle y|R\rangle|^2 |\langle R|x\rangle|^2 + |\langle y|L\rangle|^2 |\langle L|x\rangle|^2 \\ & + 2\text{Real}\left[\langle y|R\rangle\langle R|x\rangle e^{i(\alpha_R - \alpha_L)} \langle y|L\rangle^* \langle L|x\rangle^*\right] \end{aligned}$$

The observed probability, which is the result of many "identical" measurements in the laboratory, is an average over all values of the extra phases (they are random).

This involves integrating over the relative phase, i.e.,

$$\frac{1}{2\pi} \int_0^{2\pi} e^{i(\alpha_R - \alpha_L)} d(\alpha_R - \alpha_L) = 0$$

It is clear that the interference term averages to zero and we get the classical result!!

Pure States, Unpure States and Density Operators

If the photon were in the state $|x\rangle$, then we would have, for some

linear operator \hat{A} ,

$$\langle \hat{A} \rangle = \langle x | \hat{A} | x \rangle = \text{expectation or average value} \quad (77)$$

We define a property of an operator called the **trace** as

$$\text{Tr} \hat{Q} = \sum_j \langle q_j | \hat{Q} | q_j \rangle = \text{sum of diagonal matrix elements} = \sum_j (\hat{Q})_{jj} \quad (78)$$

that is, the sum over the diagonal matrix elements.

Some Properties of the Trace:

$$\begin{aligned} \text{Tr}(AB) &= \text{Tr}(BA) \\ \text{Tr}(cB) &= c\text{Tr}(B) \\ \text{Tr}(c(A+B)) &= \text{Tr}(cA) + \text{Tr}(cB) = c\text{Tr}(A) + c\text{Tr}(B) \end{aligned} \quad (79)$$

Definition : A **density operator** is a **positive, Hermitian operator** \hat{W} with a discrete eigenvalue spectrum such that, given **any** orthonormal basis set $\{|\phi_k\rangle\}$, we have

$$\text{Tr} \hat{W} = 1 = \sum_k W_{kk} = \sum_k \langle \phi_k | \hat{W} | \phi_k \rangle \quad (80)$$

where W_{kk} is the diagonal matrix element (in the basis) of the density operator \hat{W}

Quantum theory says :

- (1) A density operator **exists** for **every** real physical system (in the same way that every physical system can be represented by a state vector or ket).
- (2) The expectation value of an operator \hat{A} is given by

$$\langle \hat{A} \rangle = \text{Tr}(\hat{W}\hat{A}) \quad (81)$$

Let us choose a simple example of a density operator to get some handle on what this postulate is saying.

In particular, let us choose as our density operator \hat{W} the projection operator onto a 1-dimensional subspace spanned by some vector $|\alpha\rangle$

$$\hat{W} = |\alpha\rangle\langle\alpha| \quad (82)$$

This is an **idempotent** operator since $\hat{W}^2 = \hat{W}$ and thus has eigenvalues $w_k = 0, 1$ only, i.e.,

$$\begin{aligned}
\hat{W}|\beta\rangle &= \beta|\beta\rangle \\
\hat{W}^2|\beta\rangle &= \hat{W}\beta|\beta\rangle = \beta^2|\beta\rangle = \hat{W}|\beta\rangle = \beta|\beta\rangle \\
(\beta^2 - \beta)|\beta\rangle &= 0 \\
\beta^2 - \beta = 0 &\rightarrow \beta = 0,1
\end{aligned}$$

For equation (82) the eigenvector corresponding to eigenvalue 1 is $|\alpha\rangle$.

Properties of the Density Operator

$$\sum_k w_k = 0 + 1 = 1 = \text{Tr}\hat{W} \tag{83}$$

$$\langle a|\hat{W}|a\rangle = |\langle a|\alpha\rangle|^2 \geq 0$$

so that all required properties for a density operator are, in fact, satisfied by (82).

If we denote the eigenvalues of \hat{W} by w_k and the corresponding eigenvectors by $|w_k\rangle$ so that

$$\hat{W}|w_k\rangle = w_k|w_k\rangle$$

then, since \hat{W} has a discrete spectrum, we can write \hat{W} in terms of its eigenvalues and eigenvectors as

$$\hat{W} = \sum_k w_k |w_k\rangle\langle w_k| \tag{84}$$

Since \hat{W} is Hermitian, its eigenvectors must form an orthonormal basis where

$$\langle w_k|w_j\rangle = \delta_{kj} \tag{85}$$

We now derive some other properties of this density operator object.

The spectrum of \hat{W} is the discrete set of numbers $\{w_k\}$. We then have

$$\text{Tr}\hat{W} = 1 = \sum_j \langle w_j|\hat{W}|w_j\rangle = \sum_j \langle w_j|w_j|w_j\rangle = \sum_j w_j \langle w_j|w_j\rangle \tag{86}$$

$$\text{or} \quad \sum_j w_j = 1$$

Since \hat{W} is Hermitian, we have

$$\hat{W} = \hat{W}^\dagger$$

which implies that the eigenvalues are real numbers

$$w_k = w_k^* \tag{87}$$

Using the fact that \hat{W} is defined to be a **positive** operator, we then have

$$\begin{aligned}\langle a|\hat{W}|a\rangle &= \langle a|(\sum_k w_k|w_k\rangle\langle w_k|)|a\rangle = \sum_k w_k\langle a|w_k\rangle\langle w_k|a\rangle \\ &= \sum_k w_k|\langle a|w_k\rangle|^2 \geq 0\end{aligned}$$

for any vector $|a\rangle$. This can only be true, in general, if

$$w_k \geq 0 \quad \text{for all } k \tag{88}$$

The results

$$w_k \geq 0 \quad \text{for all } k \quad \text{and} \quad \sum_k w_k = 1$$

imply that

$$0 \leq w_k \leq 1 \tag{89}$$

Returning to the simple case of equation (80), we then have

$$\langle \hat{B} \rangle = \text{Tr} \hat{W} \hat{B} = \langle \alpha | \hat{B} | \alpha \rangle = \text{expectation value of } \hat{B} \text{ in the state } |\alpha\rangle$$

Proof :

$$|w_1\rangle = |\beta\rangle, |w_2\rangle = |\alpha\rangle, \langle \alpha | \beta \rangle = 0 \tag{90}$$

$$\begin{aligned}\text{Tr} \hat{W} \hat{B} &= \sum_{k,l=1}^2 \langle \alpha | w_k \rangle \langle w_k | \hat{B} | w_l \rangle \langle w_l | \alpha \rangle \\ &= \langle \alpha | \hat{B} | \alpha \rangle\end{aligned}$$

Since the important quantities for connection to experiment will be these expectation values, we see that the state represented by \hat{W} is equally well represented by the state vector $|\alpha\rangle$ in this simple case. The density operator and the state vector are **equivalent** ways of representing a physical system in this simple case.

The most important way of distinguishing whether a state is pure or not follows from the following property of density operators :

The density operator for a **pure state** cannot be written as a linear combination of the density operators of other states, but the density operator for a "**nonpure**" state can always be so written.

This is illustrated below with some examples.

Using the $|x\rangle, |y\rangle$ basis we have

$$\begin{aligned}\langle \hat{A} \rangle &= \langle x | \hat{A} | x \rangle = \text{Tr}(\hat{W} \hat{A}) = \langle x | \hat{W} \hat{A} | x \rangle + \langle y | \hat{W} \hat{A} | y \rangle \\ &= \langle x | \hat{W} \hat{A} | x \rangle + \langle y | \hat{W} \hat{A} | y \rangle \\ &= \langle x | \hat{W} | x \rangle \langle x | \hat{A} | x \rangle + \langle x | \hat{W} | y \rangle \langle y | \hat{A} | x \rangle + \langle y | \hat{W} | x \rangle \langle x | \hat{A} | y \rangle + \langle y | \hat{W} | y \rangle \langle y | \hat{A} | y \rangle\end{aligned}$$

This implies that

$$\langle x|\hat{W}|x\rangle = 1 \quad \text{and} \quad \langle x|\hat{W}|y\rangle = \langle y|\hat{W}|x\rangle = \langle y|\hat{W}|y\rangle = 0$$

or

$$\hat{W} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |x\rangle\langle x| \quad (91)$$

which is a linear combination of density operators and says that $|x\rangle$ is a **pure state**.

Now suppose that the photon is in the state

$$|\psi\rangle = \frac{e^{i\alpha_x}}{\sqrt{2}}|x\rangle + \frac{e^{i\alpha_y}}{\sqrt{2}}|y\rangle$$

where we know that the phases are equal, $\alpha_x = \alpha_y$ (the relative phase $\alpha_x - \alpha_y = 0$ between the components is **known exactly** in this state) so that

$$|\psi\rangle = e^{i\alpha_x} \left(\frac{1}{\sqrt{2}}|x\rangle + \frac{1}{\sqrt{2}}|y\rangle \right)$$

The terms $e^{i\alpha_x}$, etc, are called phase factors.

But, since all states but have length 1, we can ignore the overall phase factor and write

$$|\psi\rangle = \frac{1}{\sqrt{2}}|x\rangle + \frac{1}{\sqrt{2}}|y\rangle$$

This says that the probability = 1/2 that the photon behaves like $|x\rangle$ and the probability = 1/2 that it behaves like $|y\rangle$.

In this case, we have

$$\begin{aligned} \langle \hat{A} \rangle &= \langle \psi | \hat{A} | \psi \rangle = \frac{1}{2} \left[\langle x | \hat{A} | x \rangle + \langle x | \hat{A} | y \rangle + \langle y | \hat{A} | x \rangle + \langle y | \hat{A} | y \rangle \right] \\ &= Tr(\hat{W}\hat{A}) = \langle x | \hat{W}\hat{A} | x \rangle + \langle y | \hat{W}\hat{A} | y \rangle = \langle x | \hat{W}\hat{A} | x \rangle + \langle y | \hat{W}\hat{A} | y \rangle \\ &= \langle x | \hat{W} | x \rangle \langle x | \hat{A} | x \rangle + \langle x | \hat{W} | y \rangle \langle y | \hat{A} | x \rangle + \langle y | \hat{W} | x \rangle \langle x | \hat{A} | y \rangle + \langle y | \hat{W} | y \rangle \langle y | \hat{A} | y \rangle \end{aligned}$$

which implies that

$$\langle x | \hat{W} | x \rangle = \frac{1}{2} = \langle x | \hat{W} | y \rangle = \langle y | \hat{W} | x \rangle = \langle y | \hat{W} | y \rangle$$

or

$$\hat{W} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = |\psi\rangle\langle\psi| \quad (92)$$

So, again we have a pure state.

But what happens if we **only know** that the probability = 1/2 that the photon behaves like $|x\rangle$ and the probability = 1/2 that it behaves like $|y\rangle$.

This says that the state vector is

$$|\psi\rangle = a|x\rangle + b|y\rangle$$

where we **only know** that $|a|^2 = |b|^2 = \frac{1}{2}$. Let us choose

$$a = \frac{e^{i\alpha_a}}{\sqrt{2}} \quad \text{and} \quad b = \frac{e^{i\alpha_b}}{\sqrt{2}}$$

We do not have any phase information in this case. In addition, the phases values could be different in each separate experiment. This means that we must average over the relative phase $\alpha_a - \alpha_b$ when computing the probabilities and thus all interference effects will vanish.

When we calculate the expectation value we have

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle = \frac{1}{2} \left[\langle x | \hat{A} | x \rangle + e^{-i(\alpha_a - \alpha_b)} \langle x | \hat{A} | y \rangle + e^{i(\alpha_a - \alpha_b)} \langle y | \hat{A} | x \rangle + \langle y | \hat{A} | y \rangle \right]$$

and when we average over the relative phase we obtain

$$\langle \hat{A} \rangle = \frac{1}{2} \langle x | \hat{A} | x \rangle + \frac{1}{2} \langle y | \hat{A} | y \rangle$$

Again, we must have

$$\begin{aligned} \langle \hat{A} \rangle &= \text{Tr}(\hat{W}\hat{A}) = \langle x | \hat{W}\hat{A} | x \rangle + \langle y | \hat{W}\hat{A} | y \rangle = \langle x | \hat{W}\hat{A} | x \rangle + \langle y | \hat{W}\hat{A} | y \rangle \\ &= \langle x | \hat{W} | x \rangle \langle x | \hat{A} | x \rangle + \langle x | \hat{W} | y \rangle \langle y | \hat{A} | x \rangle + \langle y | \hat{W} | x \rangle \langle x | \hat{A} | y \rangle + \langle y | \hat{W} | y \rangle \langle y | \hat{A} | y \rangle \end{aligned}$$

which implies that

$$\langle x | \hat{W} | x \rangle = \frac{1}{2} = \langle y | \hat{W} | y \rangle \quad \text{and} \quad \langle y | \hat{W} | x \rangle = \langle x | \hat{W} | y \rangle = 0$$

or

$$\hat{W} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} |x\rangle\langle x| + \frac{1}{2} |y\rangle\langle y| = \text{probability}(x) |x\rangle\langle x| + \text{probability}(y) |y\rangle\langle y| \quad (93)$$

This is a **nonpure** or **mixed state**.

So, we **have a pure state only if the relative phase information is known exactly**.

Unpolarized Light

Consider the following experiment.

We have a beam of monochromatic light that is composed of photons from two sources which output photons in the states $|\psi_1\rangle$ or $|\psi_2\rangle$, respectively. The sources emit the photons randomly and are independent of each other, which implies that we cannot tell which source a particular photon comes from.

We assign these probabilities

p_1 = probability that a photon comes from source #1
 p_2 = probability that a photon comes from source #2

where $p_1 + p_2 = 1$. Now the probability that a particular observed photon transfers \hbar is

$$p_+ = p_1 \langle R | \psi_1 \rangle^2 + p_2 \langle R | \psi_2 \rangle^2$$

and the probability that it transfers $-\hbar$ is

$$p_- = p_1 \langle L | \psi_1 \rangle^2 + p_2 \langle L | \psi_2 \rangle^2$$

This implies that the average value of the angular momentum transfer for the beam of photons is

$$\begin{aligned}
 \langle \hat{J}_z \rangle &= \hbar p_+ - \hbar p_- = \hbar p_1 \langle R | \psi_1 \rangle^2 + \hbar p_2 \langle R | \psi_2 \rangle^2 - \hbar p_1 \langle L | \psi_1 \rangle^2 - \hbar p_2 \langle L | \psi_2 \rangle^2 \\
 &= p_1 \left[\hbar \langle R | \psi_1 \rangle^2 - \hbar \langle L | \psi_1 \rangle^2 \right] + p_2 \left[\hbar \langle R | \psi_2 \rangle^2 - \hbar \langle L | \psi_2 \rangle^2 \right] \\
 &= p_1 \langle \hat{J}_z \rangle + p_2 \langle \hat{J}_z \rangle
 \end{aligned}$$

or, the average value of the angular momentum transfer for the beam of photons = sum over the average value in each beam weighted by the probability that photon comes from that beam.

Let me emphasize at this point that it is important to realize that the statement

the photon is either in the state $|\psi_1\rangle$ or $|\psi_2\rangle$ but we do not know which

is **NOT** the same statement as

the photon is in a state which is a superposition of $|\psi_1\rangle$ and $|\psi_2\rangle$

In the second case, we are saying the **relative phase is known** as in the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}|x\rangle + \frac{1}{\sqrt{2}}|y\rangle$$

which we found to be a pure state.

Being in a superposition implies that we know the relative phase of the components.

In the first case, however, we are saying that the relative phase is unknown and, as we have seen, interference effects will vanish.

In pure states, we have superpositions and the probability amplitude rules apply. In nonpure or mixed states, where the system is in one of several states with definite probabilities, we find weighted averages (weighted with the state probabilities) of the value in each state. We use addition of probabilities with no interference effects,

which as we have seen, is equivalent to saying the relative phase is unknown.

Unpolarized light has equal probability of being in any polarization state. It is just a special nonpure or mixed state. No relative phase information is known for unpolarized light.

How Does the Polarization State Vector Change in Physical Systems?

Up to now we have been considering devices such as polaroids and prisms, which are **"go-nogo"** devices. Some photons get through and some do not for these devices depending on their polarization state.

We now consider devices where all the photons get through no matter what their polarization state is, but the device changes the incident polarization state in some way.

In particular, we consider the example of a **"birefringent"** crystal, such as calcite. A calcite crystal has a **preferred** direction called the **optic** axis. The crystal has a different index of refraction for light polarized parallel to the optic axis than it has for light polarized perpendicular to the optic axis. We assume that the optic axis is in the x-y plane and send a beam of photons in the z-direction. Photons polarized perpendicular to the optic axis are called **ordinary** and are in the state $|o\rangle$ and photons polarized parallel to the optic axis are called **extraordinary** and are in the state $|e\rangle$.

The set of states $\{|o\rangle, |e\rangle\}$ forms an orthonormal basis and general photon states interacting with a calcite crystal are written as superpositions of these basis states.

This is an example of a **general rule in quantum mechanics**.

If we are doing an experiment using a particular measuring device that measures the observable \hat{Q} , then we should use as the basis for all states, the eigenvectors of \hat{Q} . As we shall see, this requirement pushes us to ask the correct experimental questions (those that quantum mechanics can answer). This particular basis is called the **home** space for the experiment.

Now, as we saw earlier, the phase of a light wave with wavelength λ as it propagates through a medium in the z-direction is given by the quantity

$$\phi = e^{ikz} \tag{94}$$

with

$$k = \frac{2\pi}{\lambda} = \frac{n\omega}{c} \tag{95}$$

where n = index of refraction, $\omega = 2\pi\nu$, ν = frequency and c = speed of light.

Since the phase depends on the index of refraction, the effect of passing through a calcite crystal is to change the **relative phase** of the $|o\rangle$ and $|e\rangle$ components making up the superposition.

We assume that the state of the photon entering the calcite crystal is

$$|\psi_{in}\rangle = |e\rangle\langle e|\psi_{in}\rangle + |o\rangle\langle o|\psi_{in}\rangle \quad (96)$$

The two components have different indices of refraction n_e and n_o , respectively .

If the beam passes through a length ℓ of calcite, then the state upon leaving is given by inserting phase changes for each component and remembering that the component phases change differently.

$$|\psi_{out}\rangle = e^{ik_e\ell}|e\rangle\langle e|\psi_{in}\rangle + e^{ik_o\ell}|o\rangle\langle o|\psi_{in}\rangle = \hat{U}_\ell|\psi_{in}\rangle \quad (97)$$

where

$$\hat{U}_z = e^{ik_e z}|e\rangle\langle e| + e^{ik_o z}|o\rangle\langle o| \quad (98)$$

is a **"time development"** operator of some sort since ℓ = distance traveled in a time t .

Now we define two new quantities which will be with us throughout our study of Quantum Mechanics.

For transitions between two states (**in** and **out** in this case)

$$\langle\phi|\psi_{out}\rangle = \langle\phi|\hat{U}_\ell|\psi_{in}\rangle = \text{the transition amplitude for a photon to enter the calcite in state } |\psi_{in}\rangle \text{ and leave in state } |\phi\rangle \quad (99)$$

$$|\langle\phi|\psi_{out}\rangle|^2 = |\langle\phi|\hat{U}_\ell|\psi_{in}\rangle|^2 = \text{the transition probability} \quad (100)$$

To proceed any further, we need to find out more about \hat{U}_z . Now

$$\begin{aligned} |\psi_z\rangle &= \text{state of the photon after traveing a distance } z \text{ through calcite} \\ &= \hat{U}_z|\psi_{in}\rangle \end{aligned} \quad (101)$$

From the form of \hat{U}_z we have

$$\begin{aligned} \hat{U}_{z+\varepsilon} &= e^{ik_e(z+\varepsilon)}|e\rangle\langle e| + e^{ik_o(z+\varepsilon)}|o\rangle\langle o| \\ &= (e^{ik_e\varepsilon}|e\rangle\langle e| + e^{ik_o\varepsilon}|o\rangle\langle o|)(e^{ik_e z}|e\rangle\langle e| + e^{ik_o z}|o\rangle\langle o|) \end{aligned}$$

or

$$\hat{U}_{z+\varepsilon} = \hat{U}_\varepsilon \hat{U}_z \quad (102)$$

This is general result for all time development operators, namely,

$$\hat{U}_{t+t'} = \hat{U}_{t'} \hat{U}_t$$

This implies that

$$|\psi_{z+\varepsilon}\rangle = \hat{U}_{z+\varepsilon}|\psi_{in}\rangle = \hat{U}_\varepsilon \hat{U}_z|\psi_{in}\rangle = \hat{U}_\varepsilon|\psi_z\rangle \quad (103)$$

Now let $\varepsilon \rightarrow 0$ such that $k_o\varepsilon \ll 1$ and $k_e\varepsilon \ll 1$ and we can write (to 1st order)

$$\begin{aligned}
\hat{U}_\varepsilon &= e^{ik_e\varepsilon}|e\rangle\langle e| + e^{ik_o\varepsilon}|o\rangle\langle o| \\
&= (1 + ik_e\varepsilon)|e\rangle\langle e| + (1 + ik_o\varepsilon)|o\rangle\langle o| \\
&= \hat{I} + i\varepsilon\hat{K}
\end{aligned} \tag{104}$$

where

$$\hat{I} = |e\rangle\langle e| + |o\rangle\langle o| \quad \text{and} \quad \hat{K} = k_e|e\rangle\langle e| + k_o|o\rangle\langle o| \tag{105}$$

Now, the relation

$$\hat{K} = k_e|e\rangle\langle e| + k_o|o\rangle\langle o| \tag{106}$$

is an expansion of an operator in terms of its eigenvalues and the corresponding projection operators (eigenvectors). It says that the eigenvectors of \hat{K} are $|e\rangle$ and $|o\rangle$ with eigenvalues k_e and k_o , respectively.

This illustrates the **awesome power** in these methods!!

We then have

$$|\psi_{z+\varepsilon}\rangle = (\hat{I} + i\varepsilon\hat{K})|\psi_z\rangle \tag{107}$$

or

$$|\psi_{z+\varepsilon}\rangle - |\psi_z\rangle = i\varepsilon\hat{K}|\psi_z\rangle \tag{108}$$

or

$$\lim_{\varepsilon \rightarrow 0} \frac{|\psi_{z+\varepsilon}\rangle - |\psi_z\rangle}{\varepsilon} = i\hat{K}|\psi_z\rangle \tag{109}$$

which gives the differential equation for the **"time"** development of the state vector

$$\frac{d}{dx}|\psi_z\rangle = i\hat{K}|\psi_z\rangle \tag{110}$$

It is clearly similar to the differential equation we obtained earlier for the time development operator. If we follow the analogy, then we should have

$$\hat{K} = \text{Hermitian operator} \quad \text{and} \quad \hat{U}_z = \text{unitary operator}$$

Let us derive some important results from equation (31). We have, using the x-y basis

$$\begin{aligned}
\langle x|\psi_{z+\varepsilon}\rangle - \langle x|\psi_z\rangle &= i\varepsilon\langle x|\hat{K}|\psi_z\rangle \\
&= i\varepsilon\langle x|\hat{K}\hat{I}|\psi_z\rangle = i\varepsilon\langle x|\hat{K}|x\rangle\langle x|\psi_z\rangle + i\varepsilon\langle x|\hat{K}|y\rangle\langle y|\psi_z\rangle
\end{aligned} \tag{111}$$

or the change in the x-component of $|\psi_z\rangle$ as we move an infinitesimal amount ε has one part proportional to the x-component of $|\psi_z\rangle$ and a second part proportional to the y-component of $|\psi_z\rangle$.

Similarly, we have

$$\begin{aligned}
\langle y|\psi_{z+\varepsilon}\rangle - \langle y|\psi_z\rangle &= i\varepsilon\langle y|\hat{K}|\psi_z\rangle \\
&= i\varepsilon\langle y|\hat{K}\hat{I}|\psi_z\rangle = i\varepsilon\langle y|\hat{K}|x\rangle\langle x|\psi_z\rangle + i\varepsilon\langle y|\hat{K}|y\rangle\langle y|\psi_z\rangle
\end{aligned} \tag{112}$$

Now, since **no photons are lost** as we pass through, we must have

$$\langle \psi_{z+\varepsilon}|\psi_{z+\varepsilon}\rangle = 1 = \langle \psi_z|\psi_z\rangle \tag{113}$$

for all z . Equations (111) and (112) give

$$\begin{aligned}
\langle \psi_{z+\varepsilon}|\psi_{z+\varepsilon}\rangle &= \langle \psi_z|\psi_z\rangle + i\varepsilon\left[\langle x|\hat{K}|x\rangle - \langle x|\hat{K}|x\rangle^*\right]\langle x|\psi_z\rangle^2 \\
&\quad + i\varepsilon\left[\langle y|\hat{K}|y\rangle - \langle y|\hat{K}|y\rangle^*\right]\langle y|\psi_z\rangle^2 \\
&\quad + i\varepsilon\left[\langle x|\hat{K}|y\rangle - \langle x|\hat{K}|y\rangle^*\right]\langle y|\psi_z\rangle\langle x|\psi_z\rangle^* \\
&\quad + i\varepsilon\left[\langle y|\hat{K}|x\rangle - \langle y|\hat{K}|x\rangle^*\right]\langle x|\psi_z\rangle\langle y|\psi_z\rangle^*
\end{aligned}$$

which says that we must have

$$\begin{aligned}
\langle x|\hat{K}|x\rangle &= \langle x|\hat{K}|x\rangle^* \quad , \quad \langle y|\hat{K}|y\rangle = \langle y|\hat{K}|y\rangle^* \\
\langle x|\hat{K}|y\rangle &= \langle x|\hat{K}|y\rangle^* \quad , \quad \langle y|\hat{K}|x\rangle = \langle y|\hat{K}|x\rangle^*
\end{aligned}$$

This says that \hat{K} is a Hermitian operator.

Finally, one can show that $\hat{U}_z^+\hat{U}_z = \hat{I}$ so that \hat{U}_z is unitary as for the time transformation operator.

From our earlier discussion, we identify

$$\begin{aligned}
\hat{U}_z &= \text{transformation operator} \\
\hat{K} &= \text{generator of the transformation}
\end{aligned}$$

Calculating the Transition Probability

We defined the transition probability as

$$T(z) = \left| \langle \phi | \psi_{z,out} \rangle \right|^2 = \left| \langle \phi | \hat{U}_z | \psi_{in} \rangle \right|^2 \tag{114}$$

Using $\hat{U}_z = e^{ik_e z} |e\rangle\langle e| + e^{ik_o z} |o\rangle\langle o|$ and $|\psi_{in}\rangle = a|o\rangle + b|e\rangle$, where $|a|^2 + |b|^2 = 1$, we get

$$\begin{aligned}
T(z) &= \left| \langle \phi | \left(e^{ik_e z} |e\rangle\langle e| + e^{ik_o z} |o\rangle\langle o| \right) (a|o\rangle + b|e\rangle) \right|^2 \\
&= \left| \langle \phi | \left(be^{ik_e z} |e\rangle + ae^{ik_o z} |o\rangle \right) \right|^2 = \left| be^{ik_e z} \langle \phi | e \rangle + ae^{ik_o z} \langle \phi | o \rangle \right|^2
\end{aligned} \tag{115}$$

Now let us ask a **specific** question.

Suppose $a = \frac{1}{\sqrt{2}} = -ib$, which means the that photon entering the calcite crystal is an LCP photon.

What is the probability that it will exit as a RCP photon? This means we choose

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|o\rangle + i|e\rangle)$$

or

$$\langle\phi|e\rangle = \frac{i}{\sqrt{2}} \quad \text{and} \quad \langle\phi|o\rangle = \frac{1}{\sqrt{2}}$$

We then get

$$\begin{aligned} T(z) &= \left| be^{ik_e z} \langle\phi|e\rangle + ae^{ik_o z} \langle\phi|o\rangle \right|^2 = \left| \frac{i}{\sqrt{2}} e^{ik_e z} \frac{i}{\sqrt{2}} + \frac{1}{\sqrt{2}} e^{ik_o z} \frac{1}{\sqrt{2}} \right|^2 \\ &= \frac{1}{4} \left| e^{ik_o z} - e^{ik_e z} \right|^2 = \frac{1}{4} \left(1 + 1 - e^{i(k_o - k_e)z} - e^{-i(k_o - k_e)z} \right) \\ &= \frac{1}{2} (1 - \cos(k_o - k_e)z) \end{aligned}$$

If we choose $(k_o - k_e)z = \pi$, then $T=1$ and **all** the LCP photons are turned into RCP photons by a calcite crystal of **just the right length**.

This simple example clearly exhibits the power of these techniques.