For the adventurous students……

Polarization

Now let us use all formalism that we have developed to completely understand a real-world property associated with photons(or light waves), namely, **polarization**.

Before starting out, we need to digress and learn/review a few more mathematical things about matrices and better understand the quantum mechanics we have developed.

More about Matrix Operators

As have seen, operators are very important objects in quantum theory; some operators directly represent physical properties and others transform states.

We now expand on the earlier discussion of ways of writing operators with some additional information about matrices.

A matrix is array of numbers, such as $\begin{pmatrix} 1 & 4 \\ 7 & 9 \end{pmatrix}$ which is two by two or 2×2 matrix or by $\begin{pmatrix} 1 & 34 \\ 9 & 18 \\ 5 & 12 \end{pmatrix}$ which is three by two or 3 × 2 matrix.

A matrix can have any number of rows and columns, including one by something or something by one.

Each number in matrix is called a **matrix element**.

Of course, we can always use letters to represent elements when we are not sure what the number is, or when we deliberately do not want to refer to a specific number.

To make such arrays useful, mathematicians invented series of rules that allow us to add, subtract, and multiply matrices together.

You can add and subtract matrices if they have same number of rows and columns —> add elements.

$$\begin{pmatrix} 1 & 4 \\ 7 & 9 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 12 & 10 \end{pmatrix}$$

We will not have to add or subtract matrices; we will only need to multiply them together.

We can multiply two matrices if number of columns in one is equal to the number of rows in another. For example

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} aA + bC & aB + bD \\ cA + dC & cB + dD \end{pmatrix}$$

which looks fiendishly complicated, but its actually not so really complicated.

What we are doing is multiplying a row in first matrix by a column in second.

Look at top left element of resulting matrix, aA+bC.

We can see that we have combined elements from top row of first matrix and those from left column of second.

Similarly, top right element, aB + bD, is built from top row of first matrix and right column of second.

Work your way through the elements and you will see how it is all put together.

Quantum states can be written in matrix form(remember color/hardness discussion).

For example, the eigenstates of vertical Stern-Gerlach(S-G) operator $\hat{S}_z$ can be written as

$$|U\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad |D\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \textbf{in UP/DOWN basis}$$

Think of the top rows as containing the amplitude for a particle to be in state |U⟩ and bottom row as the amplitude for it to be in state |D⟩.

In formal sense, this says that state can be thought of as array of amplitudes for transitions to other states (normally eigenstates).

States such as |R⟩ and |L⟩ are written in the UP/DOWN basis as

$$|R\rangle = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad |L\rangle = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

since, for example, amplitude for |R⟩ to transition to |U⟩ is $1/\sqrt{2}$ and so on.

Now we can also write the vertical S-G operator $\hat{S}_z$ in the form of a matrix

$$\hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and apply it to states and see what happens:

$$\hat{S}_z \left|U\right\rangle = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1\times 1 + 0\times 0 \\ 0\times 1 + (-1)\times 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2}\left|U\right\rangle$$

which is exactly what we wanted to happen(they are eigenstates!).

Similarly, we can show that

$$\hat{S}_z \left|D\right\rangle = -\frac{\hbar}{2}\left|D\right\rangle$$

—> |U⟩, |D⟩ = eigenstates of $\hat{S}_z$

For completeness, we point out that other two S-G operators $\hat{S}_x$ and $\hat{S}_y$ can also be written in the form of matrices.

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Although any operator can be written in matrix form, this is not the same as saying the operator **is** a matrix.

It turns out that there is more than one way to represent an operator mathematically and the matrix form is only one example.

The first step in converting an operator into a matrix is to choose the basis states to be employed.

A given operator will have different matrix representations, depending on basis set being used.

To show how this works, we derive the $\hat{S}_z$ operator using the (UP,DOWN) basis.

We should get result above. To represent $\hat{S}_z$ as matrix, such as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

we need some way of calculating various elements from this representation.

Before we can do this, we have to find a way of writing ⟨U| and ⟨D| in matrix form.

Consider the basic "braket" relationship ⟨U | U⟩ = 1= number.

The rules of matrix multiplication (row x column) —> the only way to construct something like ⟨U | ψ⟩ in matrix form, given that

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$$

is to write

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (1 \times a + 0 \times b) = a$$ as expected.

From this result, we must then have

$$\langle U| = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad , \quad \langle D| = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

In fact, for any state

$$|\phi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$$

we can write

$$\langle\phi| = \begin{pmatrix} a^* & b^* \end{pmatrix}$$

Sticking with UP and DOWN, observe the simple calculation shown below

$$\langle U|\,\hat{S}_z\,|U\rangle = \langle U|\,\frac{\hbar}{2}\,|U\rangle = \frac{\hbar}{2}\,\langle U\mid U\rangle = \frac{\hbar}{2}$$

or in matrix form

$$\frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \frac{\hbar}{2}$$

The simple calculation tells us something more general.

Having pinned down the top left matrix element,

$$\begin{pmatrix} \langle U|\,\hat{S}_z\,|U\rangle & b \\ c & d \end{pmatrix}$$

You can, by now, guess what the overall pattern is going to be

But to be sure, try it with one of other elements.

Construct

$$\langle D|\,\hat{S}_z\,|U\rangle = \langle D|\,\frac{\hbar}{2}\,|U\rangle = \frac{\hbar}{2}\,\langle D\mid U\rangle = \frac{\hbar}{2}(0) = 0$$

and then do same calculation in matrix form:

$$\begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 1\times 1+0\times 0 \\ 0\times 1+(-1)\times 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

Thus, we have found another element

$$\begin{pmatrix} \langle U|\,\hat{S}_z\,|U\rangle & b \\ \langle D|\,\hat{S}_z\,|U\rangle & d \end{pmatrix}$$

In fact, we can easily fill in the whole matrix

$$\begin{pmatrix} \langle U|\,\hat{S}_z\,|U\rangle & \langle U|\,\hat{S}_z\,|D\rangle \\ \langle D|\,\hat{S}_z\,|U\rangle & \langle D|\,\hat{S}_z\,|D\rangle \end{pmatrix}$$

as discussed earlier in class.

The generalization to any other operator is straightforward.

If have some operator $\hat{O}$ and a basis {|i⟩}, then matrix representing the operator is written as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & . & . \\ . & a_{22} & a_{23} & . & \\ . & . & a_{33} & . & . \\ a_{14} & . & . & a_{44} & . \\ . & . & a_{53} & . & . \end{pmatrix}$$

where have filled in only some of blanks so we can see the pattern.

Matrix is always square, and number of rows and columns depends on how many states in the basis.

The matrix elements, such as a_{11}, a_{12}, and so on, are calculated from the following rule:

RULE 8: Matrix elements of operator calculated from

$$a_{11} = \langle 1|\hat{O}|1\rangle \quad , \quad a_{12} = \langle 1|\hat{O}|2\rangle \quad , \quad a_{ij} = \langle i|\hat{O}|j\rangle$$

In fact, physicists call a combination such as ⟨i|$\hat{O}$|j⟩ a matrix element.

This method of representing operator as a matrix works for any operator, but in certain situations a matrix may not be most convenient form for dealing with an operator.

In the case of position and momentum operators, for example, the appropriate bases to use have an infinite number of basis states as mentioned earlier.

As a result, there will be an infinite number of elements in the matrix -> not very useful.

One final point worth making.

An operator can be put into matrix form using any basis and thus, a given operator will have, as said earlier, different matrix representations depending on basis chosen.

However, some choices of basis are more sensible(less calculations) than others.

What happens, for example, when the basis chosen is formed from the eigenstates of the operator concerned?

Each matrix element has the form $\langle i| \hat{O} |j\rangle$ and if $|i\rangle$ and $|j\rangle$ are eigenstates, then $\hat{O}|j\rangle = e_j |j\rangle$, where e_j is eigenvalue of $\hat{O}$ for state $|j\rangle$.

Matrix element then becomes $e_j\langle i | j\rangle$.

Since $|i\rangle$ and $|j\rangle$ are a basis, we have that $\langle i|j\rangle = 0$ unless i = j.

Upshot of this discussion is that matrix is diagonal:

every matrix element is zero unless on the diagonal, and elements on the diagonal are the eigenvalues of operator acting on basis made of its eigenstates:

$$\begin{pmatrix} e_1 & 0 & 0 & 0 & . \\ 0 & e_2 & 0 & 0 & . \\ 0 & 0 & e_3 & 0 & . \\ 0 & 0 & 0 & e_4 & . \\ . & . & 0 & . & . \end{pmatrix}$$

That is enough about matrices for now.

First Thoughts about Polarization

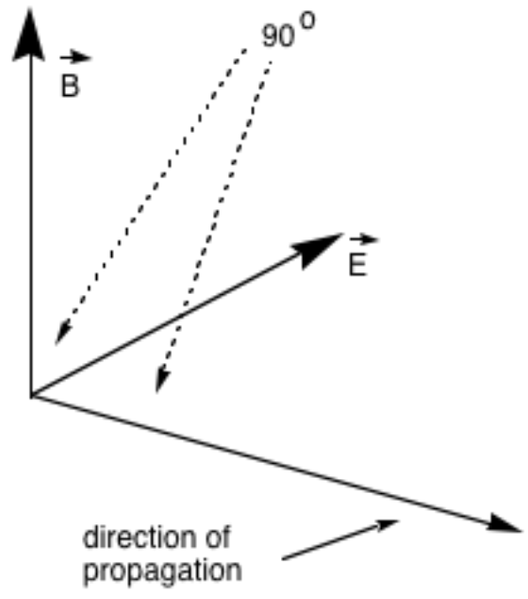
Make first pass at photon polarization in this section and return to complete theory later on.

Here we introduce polarization and some of its properties

Then we will discuss some experiments that illustrate problems with the classical description and the strangeness of quantum phenomena.

Light waves are oscillating $\vec{E}$ (electric) and $\vec{B}$ (magnetic) field vectors.

In particular, these vectors oscillate in time in a plane (a two-dimensional space) perpendicular to direction of propagation (motion) of light (—> transverse directions) as shown.



This is why you are able to see light.

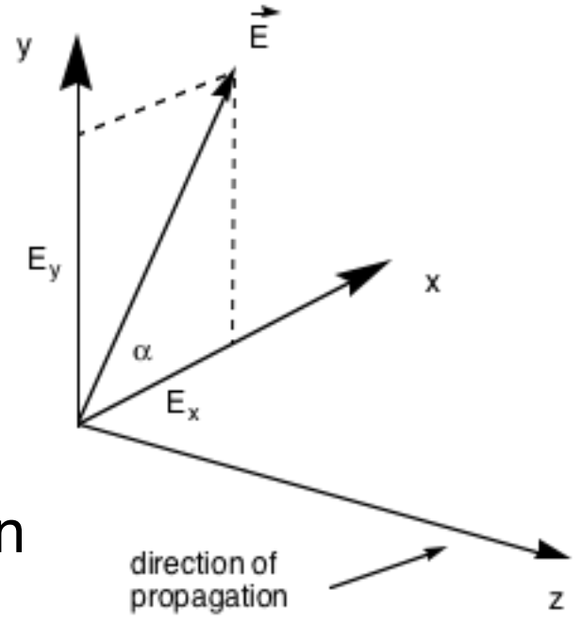
Your eye actually responding to these oscillating electromagnetic fields.

We note that the oscillations are of very-high frequency (10^{15} sec^{-1}).

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 $\vec{E}$ 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 aerials of portable radios or FM radios or TV rabbit ears to get best signal…

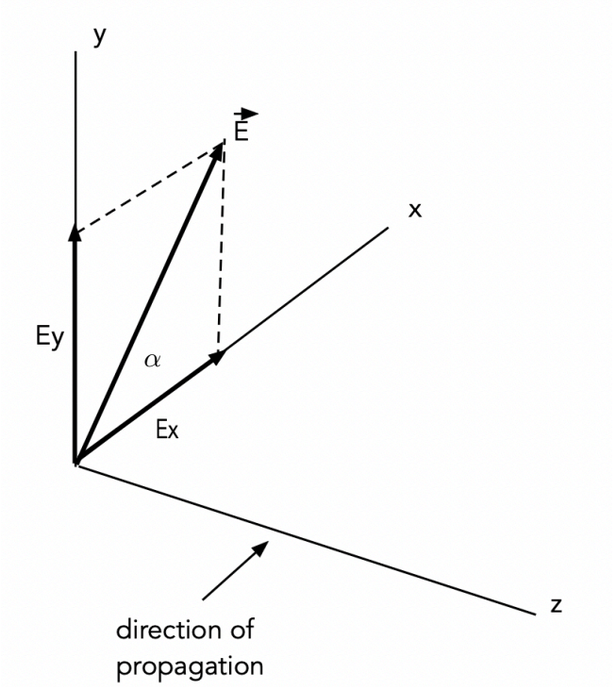
You were lining them up to be parallel or perpendicular to the polarization of electromagnetic waves being received in each case.

Also, as we will investigate in detail shortly, you have used polaroid sunglasses.

Polaroid sunglasses only allow light with polarization in single direction to get through, cutting out the light polarized in the perpendicular direction, and thus, cutting down(1/2) the total intensity(amount) of light reaching your eyes.

Now one can always choose plane containing electric and magnetic fields to be x–y plane and thus represent any vector as sum of two vectors, one in x–direction and other in y–direction, such that $\vec{E}$ vector makes angle α with x–axis as shown above.

We then say that the $\vec{E}$ vector is a combination(sum) of E_x–vector(field) and E_y–vector(field) as shown:



In terms of polarization, we say that the direction of polarization represented by $\vec{E}$ vector is (linear) combination of two polarizations, one in x–direction and one in y–direction,

i.e., polarization state of system is a linear combination of an x–polarized state and a y–polarized state.

Many materials have the property that they will only let light pass through them if polarization is in particular direction, i.e., the light is in a particular or definite polarization state.

We call this axis the **special** or **preferred** or **optic** axis of the material.

To get through light must be polarized (have the direction of polarization)

in the direction parallel to special axis.

As said earlier, you have all seen such material in sunglasses or sheets of Polaroid film.

Consider some experiments involving polaroid materials, a block of calcite and a laser.

Laser to be used produces **unpolarized** light – what does that mean? **[Demos]**

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 any directions, i.e., the polarization vector is equally likely to point in any direction.

Experimental Checks

First, send laser beam into polaroid.

Only observation is that the brightness or intensity decreases (if carefully measured it would drop by about 1/2).

If we rotate polaroid we observe no change, which says that amount of light in beam with polarization parallel to preferred direction is same no matter how we orient polaroid (how we point the preferred direction).

Fact that we get same intensity (1/2 of total) no matter what direction we choose says that light coming out of laser has equal amounts of polarization in two orthogonal directions (no matter which you two you choose!)

—> **unpolarized light** (a more formal definition later).

Note: the statement that 1/2 of light gets through polaroid when oriented in ANY direction just corresponds to fact that we can use any basis to describe physical system; all all equivalent!

When we choose a preferred direction for our world

- this is done when we pick up polaroid and orient it -

then polarization of any system we are investigating must be thought of as either being parallel or perpendicular to that chosen direction.

Setting the context, sets direction of components!!

In this case, the light from the laser now operates in the chosen context for the experiment.

Think about that!

no other cases will occur in the real world in the context(orientation of the polaroid) of the experiment

Again this would be true for any direction we might choose.

After passing through the polaroid, we say the light is polarized in a particular direction

i.e., parallel to polaroid preferred axis.

All light with polarization in orthogonal direction has been removed

That is what a polaroid does!

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

The answer must be none!

So add another polaroid and rotate it so that no light comes through.

The two polaroids are now oriented at right angles with respect to each other

and since these are only two possible polarizations with respect to chosen polaroid directions,

no light comes out.

If we rotate the pair of polaroids, keeping their relative orientation fixed

it remains the case that no light comes through.

Means, there are only two polarizations with respect to the new orientations also

and they have cancelled out also.

In the experiment, we choose the preferred direction by bringing a polaroid over - by setting experimental context.

At that point, light can be thought of as being partly made up of light polarized parallel to preferred direction and partly of light polarized perpendicular to preferred direction

and the truly amazing result, as we will see in later discussions,

will be that system does not decide what basis to use until I bring over polaroid

(**until I set the context!**).

Thus, the act of inserting the polaroid sets the context of experiment!

Think about these statements carefully!

They reflect the strange sounding explanations that result

when we try to use **words** to describe what is happening.

Mathematically, as we will see later, there will be no confusion about what is happening.

Second, send a laser beam into calcite crystal and we get two beams.

Calcite has an intrinsic(built-in) preferred direction.

Calcite causes bending of light beams where the amount of bending depends on whether 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 calcite crystal, we now see two beams coming out of the calcite.

If rotate calcite what will find?

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

Thus, the quantum world seems not to be set until I walk over

and make decision and fix orientation of calcite crystal

- until I set the context.

Again, think about statement carefully!!!

Reflect on the strange nature of the explanation when using words.

One wonders whether using words can make any sense of these phenomena!

The physical world seems to be waiting for the observer to make a decision about what measurement will be made

- waiting for context to be set.

This strange idea, which makes no sense in the classical world,

is going to get us into lot of very hot water later in our discussions.

The result is connected to fact that the act of measuring

or finding out some property of system

in some way determines or actualizes state of system being measured.

Now back to calcite beams.

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

Using 2 polaroids, where first cuts out some light and second cuts out some more,

if rule above is correct and we orient them so their preferred axes are at right angles,

then all light should be stopped.

Note that mirrors have no effect on polarization....just redirect polarized beams.

Third, now remove calcite crystal and rotate pair of polaroids until have no light intensity.

Then rotate second polaroid 45°.

Light intensity returns (actually 1/2 of 0°(or parallel) intensity).

Let us see how classical physics explains these polaroid experiments for intense(very large value of photons/sec) beams of light.

Classically, if incident beam of light polarized parallel to special or optic axis, then all of its energy gets through polaroid.

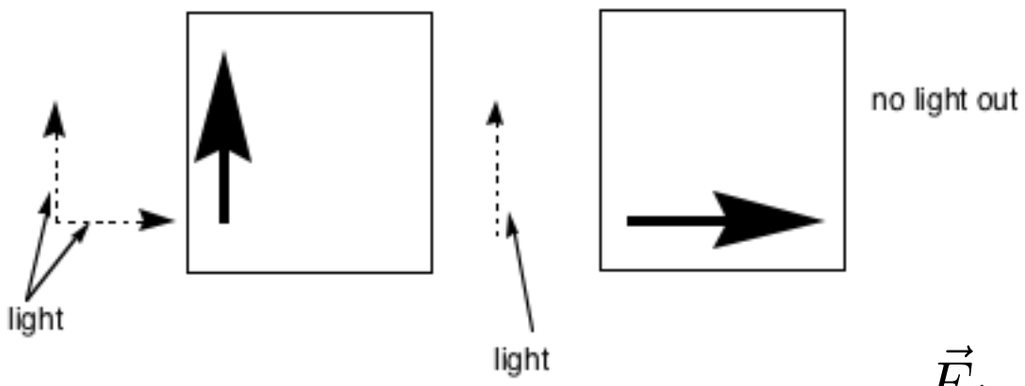
If polarized orthogonal to optic axis, then none of its energy will get through polaroid.

If polarized at angle α to preferred axis, then fraction $\cos^2\alpha$ of its energy gets through polaroid.

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

Consider 2 polaroids at right angles → intensity-out = 0.

In pictures have,

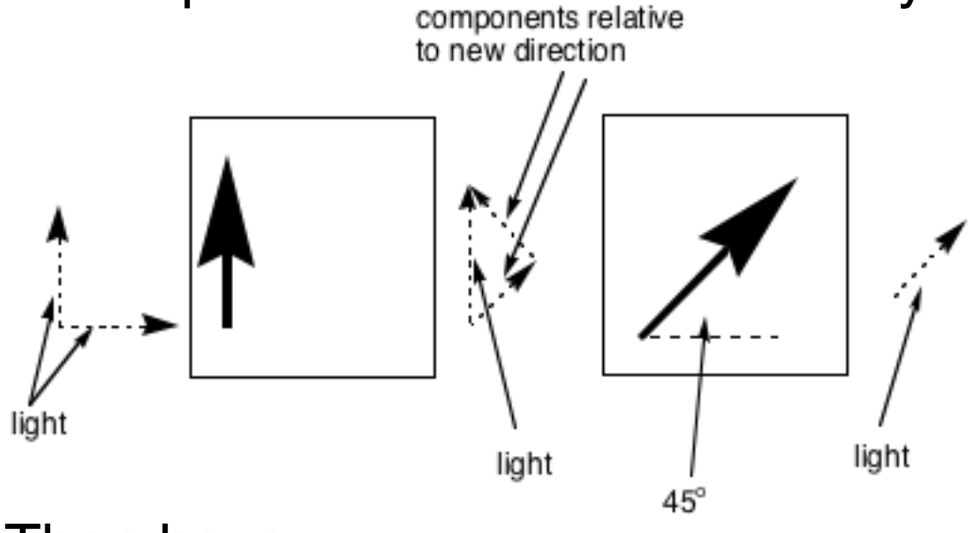


where dashed arrows represent electric field components or polarizations and solid arrows represent preferred directions of polaroids.

With components of vectors we have

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

Now 2 polaroids at 45° → intensity-out ≠ 0 as shown



With components of vectors we can derive following:

We can write unit vector that makes angle θ with x–axis as

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

Then have

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

$$\vec{E}_{intermediate} = (\vec{E}_{in}\cdot\hat{e}_y)\hat{e}_y = E_y\hat{e}_y$$

since that is what polaroid does! Finally we get

$$\vec{E}_{out} = (\vec{E}_{intermediate}\cdot\hat{n}_\theta)\hat{n}_\theta$$

$$= ((E_y\hat{e}_y)\cdot(\cos\theta\hat{e}_x + \sin\theta\hat{e}_y))(\cos\theta\hat{e}_x + \sin\theta\hat{e}_y)$$

$$= E_y\sin\theta(\cos\theta\hat{e}_x + \sin\theta\hat{e}_y)$$

Some special cases are:

$\theta = 0^\circ \rightarrow$ orthogonal polaroids $\qquad \vec{E}_{out} = 0$

$\theta = 90^\circ \rightarrow$ parallel polaroids $\qquad \vec{E}_{out} = E_y$

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

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

When beam is intense —> large numbers of photons, then it seems that a beam behaves **as if** had wave properties.

Think of analogy: water molecules making up water waves.

Remember, however, light is really composed of individual particles called photons.

Fourth, now let us add another polaroid to an earlier experiment:

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

[2 polaroids alone at 90 degrees] ⟶ light all gone

—> first polaroid reduces beams to only one direction which cannot get through second polaroid (wrong kind ….. has wrong direction).

Now that things are making some sense, we disturb system with another measurement.

- if leave system alone NO light observed

- if add third polaroid (same orientation as first), then get same result(no change); should not be any change because are not gaining any new information by this measurement .... already know that half is polarized in each direction.

- if add third polaroid (same orientation as second), then get same result(no change); again should not be any change since are still not gaining any new information ….. already know that remaining half are polarized in other direction.

- if add third polaroid oriented in different direction → different result ... Now some light gets through 2nd polaroid .....

What does this mean?

Does it mean that that the system has somehow recreated other the kind of light?

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

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

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

Let me illustrate a dilemma.

Start with large number of photons (10^{23}).

Since have a very intense beam —> it behaves like wave.

Classical physics should then be able to explain what happens.

Place polaroid in laser beam.

Half of light gets through.

Place second polaroid at right angles.

No light gets through.

Place third polaroid in between at 45°.

Half of light gets through.

Easy to explain with waves or electric field vectors and vector components as saw earlier.

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

Polaroid 2 in 45° direction (new y′ direction or $\theta = 45^\circ$)

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

Have experiment shown:

$$\vec{E}_0 \;|\; \vec{E}_1 \;|\; \vec{E}_2 \;|\; \vec{E}_3$$
$$\uparrow \qquad \uparrow \qquad \uparrow$$
$$\#1 \quad \#2 \quad \#3 \quad \text{polaroids}$$

Analyzing with vectors have:

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

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

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

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

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

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

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

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

After 2 $\quad \vec{E}_2 = \frac{E}{\sqrt{2}} \sin 45^\circ (\cos 45^\circ \hat{e}_x + \sin 45^\circ \hat{e}_y) = \frac{E}{2\sqrt{2}} (\hat{e}_x + \hat{e}_y)$

$\rightarrow \text{ energy } = \frac{E^2}{4} = \frac{1}{4} \text{ of original energy}$

After 3 $\quad \vec{E}_3 = \frac{E}{2\sqrt{2}} \hat{e}_x \rightarrow \text{ energy } = \frac{E^2}{8} = \frac{1}{8} \text{ of original energy}$ —> get some light out !

Now remove the 45 degree polaroid.... we then have

Polaroid 1 in y–direction (θ = 90° direction)

Polaroid 3 in x–direction (θ = 0° direction)

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

After 3 $\quad \vec{E}_3 = 0 \rightarrow \text{ energy } = 0$ —> get no light out!

Again, classical physics has no trouble explaining what happening for intense beams where we can talk about **dividing beam energy** between different processes.

At each step, explanation —> **fraction** of photons or **fraction** of energy does not pass though particular polaroid

and at each stage intensity of beam, related to number of photons or energy, changes in expected manner.

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

Suppose we **follow each photon** as it passes through apparatus.

Now answer not clear and fundamental dilemma of subatomic world rears its ugly head.

Remember, each individual photon cannot be subdivided - its energy cannot be split!!

As has become clear during our discussions of quantum mechanics,

this question about what will happen to particular photon under certain conditions

is not very precise and if 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 results of experiments have real significance in physics and it is only such questions that theoretical physics has to consider.

Questions and subsequent experiments devised to answer questions must be clear and precise, however.

In this case, we can make the question clear

by doing 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 polaroid.

The most important result

is that this single photon

either passes through the polaroid

or it does not.

Nothing else happens.

We never observe 1/2 of energy of single photon passing through polaroid.

Always observe either no energy or all energy.

One never observes part of photon passing through

and part getting absorbed in polaroid.

In addition, if photon gets through,

then observation shows that its polarization

is in direction parallel to optic axis of particular polaroid

instead of some arbitrary angle with respect to that axis

as it had been before it encountered polaroid.

This happens no matter what the polarization was before reaching the polaroid.

No matter what initial polarization, definite fraction of photons seem to get through polaroid.

In a beam of N identical photons, each will behave independently as the single photon did.

No experiment can determine which photon will pass through and which will not,

even though they are all identical. (sound familiar)

In each experiment, however,

exactly 1/2 of total energy

and 1/2 of 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
for passing through.

We are forced into this probabilistic point of view
by the fact that the energy of photons in an electromagnetic wave
is quantized (not continuous).

Thus, we get same result as in other experiments have discussed.

Quantization forces us to use probability!

We have managed to preserve the indivisibility of photons
or ultimately the quantization of their energy.

We are 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 will behave same way in identical experiments (within experimental error).

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

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

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

It is what is called a GO-NOGO experiment.

Does photon go through or is it absorbed?

That is only legitimate question we can ask in this case.

So if we arrange the experiment so that only one photon is inside apparatus at any one time

we have a problem, as we shall see.

Now we redo experiment with two polaroids oriented at 45°.

What happens as photon comes to first polaroid?

Either it gets through or does not —> GO-NOGO.

At second polaroid(if got through first) is again GO-NOGO.

In fact, at end of chain, either gets through or does not!

Now send another single photon through.

It is identical to first and also GO-NOGO.

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

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

What happens after large number of unpolarized photons are sent through single polaroid?

Answer is that exactly 1/2 get through!

Just look at two polaroids in experiment.

If photon gets through first polaroid, what happens at second? —> GO-NOGO

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

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

Do not know and cannot predict what any particular photon will do, but can predict what a large number will do!

That is how probability works

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 second 45° polaroid.

As have said earlier, this photon state is one of latency or propensities waiting to be actualized.

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

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 must be regarded as outside domain of quantum theory

and possibly all of physics

and cannot be relevant to any new theory we might develop.

What will quantum mechanics say about state of single photon?

A photon polarized at angle to optic axis is in very special kind of state

that we call a superposition of being polarized perpendicular to optic axis

and of being polarized parallel to optic axis.

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

The meaning of the word superposition follows from mathematical formalism and the language that we have developed.

It represents a new physical connection to mathematics.

This is suggested by any attempt to express the meaning of superposition in ordinary language(words).

If we attempt to explain behavior of photon polarized at angle to optic axis using ordinary language, then 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 logical possibilities allowed by ordinary words.

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

Its physical content is, however, precise and clear in the new mathematical formalism.

When photon encounters polaroid, we are observing it.

We are observing whether it is polarized perpendicular or parallel to optic axis of polaroid.

Effect of measurement —> end up with photon having one or other polarization.

It always makes “jump” from state of superposition to the state of definite polarization.

Which of two states it “jumps” to cannot be predicted.

Can, however, predict probability of each for large set of identical measurements.

If it "jumps" into parallel state, it passes through.

If it “jumps” into perpendicular state, it gets absorbed.

We will have a lot to say about the two new words, superposition and “jump”, as we proceed.

Will also have to show that it did not have the property that is observed after polaroid before it entered polaroid, i.e., we are not just finding out what that prior property was!!

Using Photons and Polarization to Explain How Quantum Mechanics Works?

Now look at photons and polarization in more detail (repeating much of what just said) and using the mathematical language to understand how quantum mechanics works.

As said, many experiments indicate that electromagnetic waves have the vector property called polarization.

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

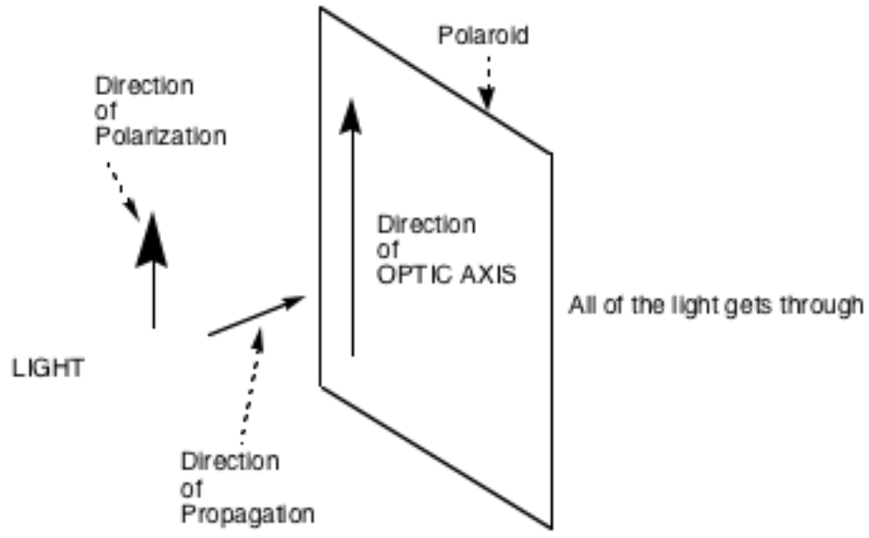
Remembering earlier discussion,

polaroid material has property that it only allows light

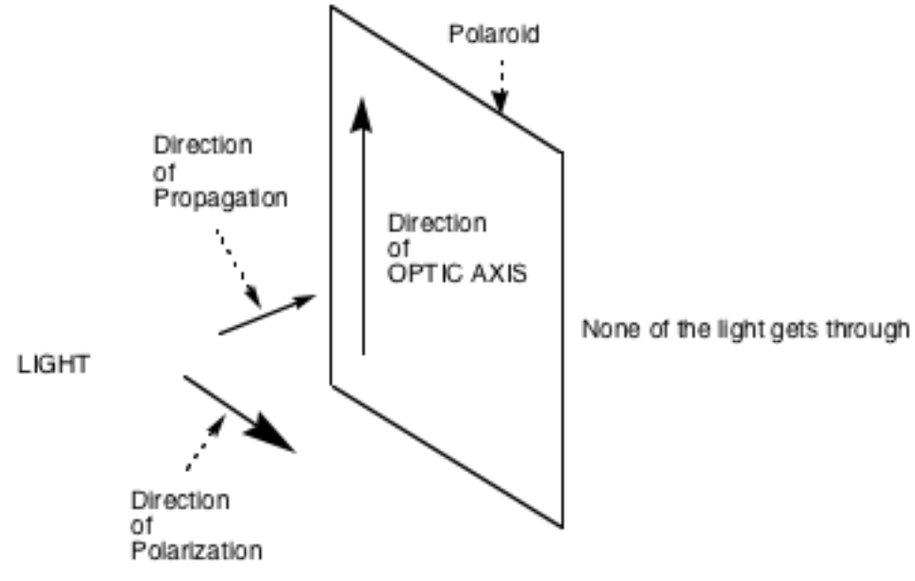
with polarization vector oriented parallel to preferred direction in polaroid(called optic axis)

to pass through material.

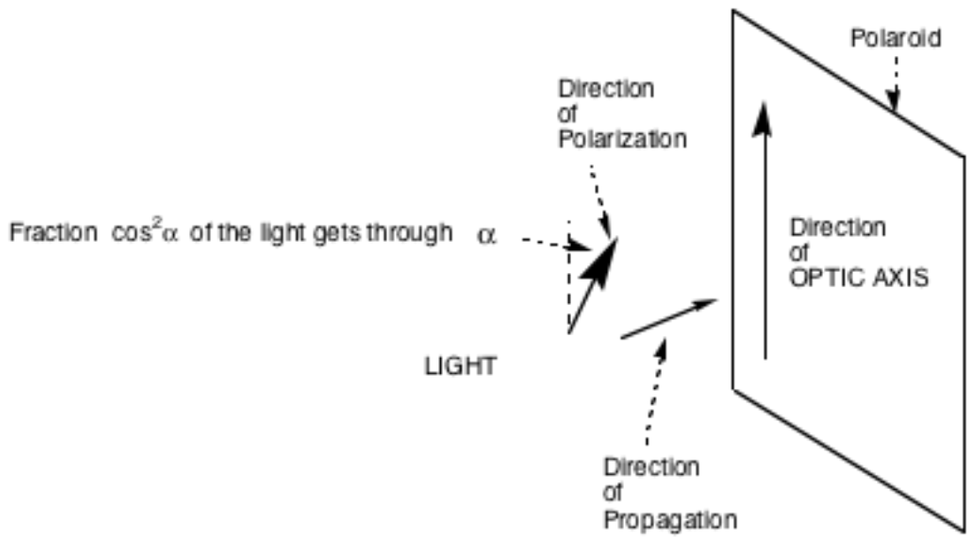
Thinking classically once again, if incident beam of light polarized parallel to optic axis (as in figure below), then experiment says that **all** of its energy gets through polaroid.



If, instead, light is polarized perpendicular to optic axis, then experiment says that **none** of its energy gets through polaroid



In more general case, if polarized at angle α to optic axis (as in figure), then experiment says that **fraction** of its energy gets through polaroid.



By definition, when we specify polarization of light, we are actually giving the direction of electric field vector associated with the light.

Polarization property or polarization vector of light depends only on the direction of the vector.

Classically, in Maxwell's theory, light waves are represented by plane electromagnetic waves

—> associated electric field vector $\vec{E}$ and associated magnetic field vector $\vec{B}$ both perpendicular to direction of propagation specified by third vector $\vec{k}$.

According to Maxwell theory, if choose(arbitrary) direction of propagation to be z–axis, specified by unit vector $\hat{\varepsilon}_z$, then $\vec{E}$ and $\vec{B}$ lie somewhere in the x–y plane, which is plane perpendicular to direction of propagation.

$\vec{E}$ and $\vec{B}$ are perpendicular to each other.

Now, any vector in 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 is called the **basis polarization vectors**.

Two standard sets of orthonormal vectors are often chosen when one discusses polarization.

One of two sets is

$$\hat{\varepsilon}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \hat{\varepsilon}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

As will see, these correspond to so-called **plane-polarized waves**.

The second orthonormal set is

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

This will correspond to -so-called **circularly-polarized waves**.

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

Plane-polarized basis:

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

Circular-polarized basis:

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

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

$$E_x(\vec{r},t) = E_x^0 e^{i(kz-\omega t+\alpha_x)}$$
$$E_y(\vec{r},t) = E_y^0 e^{i(kz-\omega t+\alpha_y)}$$

where $k = 2\pi/\lambda$, λ is wavelength, ω is angular frequency, $kz - \omega t + \alpha_x$ and $kz - \omega t + \alpha_y$ are respective **phases**, α_x and α_y are corresponding phases at $x = t = 0$ and E_x^0 and E_y^0 are (real) **amplitudes** of electric field components.

The actual(physical) electric field components are given by the real parts of the complex exponential expressions

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

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

These relations say that in an ideal 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 direction of propagation(z-direction).

For simplicity, look at the position z = 0. At that point we have

$$E_x(\vec{r},t) = E_x^0 \cos(\omega t + \alpha_x) \quad , \quad E_y(\vec{r},t) = E_y^0 \cos(\omega t + \alpha_y)$$

where

$$\omega = 2\pi f \qquad f = frequency \qquad \lambda = wavelength \qquad c = speed = \lambda f$$

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

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

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

Case #1: Orthogonal oscillations are initially in phase, i.e., αx = αy = 0 for simplicity. Then have

$$\vec{E}(\vec{r},t) = \left(E_x^0\hat{\varepsilon}_x + E_y^0\hat{\varepsilon}_y\right)\cos(\omega t) = \vec{E}_0\cos(\omega t)$$

This says that total electric field vector(its tip) oscillates with same frequency in single direction.

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

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

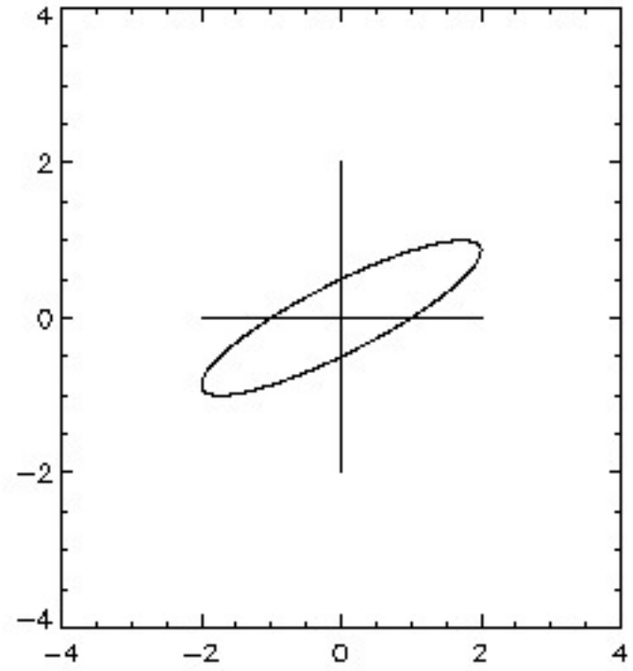
$$\vec{E}(\vec{r},t) = E_x(\vec{r},t)\hat{\varepsilon}_x + E_y(\vec{r},t)\hat{\varepsilon}_y = E_x^0\cos(\omega t)\,\hat{\varepsilon}_x + E_y^0\cos(\omega t + \alpha)\,\hat{\varepsilon}_y$$

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

This corresponds to the equation for a vector tip given by

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

which looks like (motion of tip)



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

Called **elliptically-polarized light**.

If ellipse is circle, then called **circularly-polarized light**.

If tip of electric field vector, when looked at it as the light comes straight toward us,

goes around in counterclockwise direction —> light is **right-hand circularly polarized**.

If it goes clockwise, light is **left-hand circularly polarized**.

Mathematically, the relationship between a polarization state of light and the $\vec{E}$ vector is clearly shown by the few examples below.

(1) If $E_y = 0$, then wave is plane-polarized in x-direction $\vec{E} = E_x \hat{\varepsilon}_x = E_x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

(2) If $E_x = 0$, then wave plane-polarized in y-direction $\vec{E} = E_y \hat{\varepsilon}_y = E_y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

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

$$\vec{E} = E_x \hat{\varepsilon}_x + E_y \hat{\varepsilon}_y = E_x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + E_y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = E_x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

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

$$\vec{E} = E_x \hat{\varepsilon}_R = E_x \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$$

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

$$\vec{E} = E_x \hat{\varepsilon}_L = E_x \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}$$

This set of polarization properties is almost sufficient for our discussions.

The final case to consider —> **unpolarized light**.

If initial relative x– and y–phase, i.e., $\alpha_x - \alpha_y$ not kept fixed, then the electric field vector oscillates in varying directions —> the polarization is constantly changing.

If the polarization direction changes more rapidly than we can detect, light is called **unpolarized**.

Now we return to the question - **what is a polaroid?**

In order to understand the answer to question, must first discuss **birefringence**.

An interesting effect of polarization —> substances where index of refraction of light is different for light that is linearly polarized in different directions relative to the preferred axes of material.

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

What happens when oscillating electric field passes through material?

Suppose because of structure of material, electrons in material respond more easily to oscillations in direction parallel to long axes of molecules than would respond if electric field tries to push them perpendicular to long axes —> expect different response for polarization in different directions.

The direction parallel to long axes called the optic axis.

When polarization is parallel to optic axis or along long axes of molecules, index of refraction is different then when polarization perpendicular to optic axis —> material is called **birefringent**.

It has two indices of refraction depending on direction of polarization(relative to optic axis) inside the substance.

Index of refraction affects electric field as propagates through material as shown

$$\vec{E}(\vec{r},t) = \vec{E}_0 \cos(kz - \omega t) = \vec{E}_0 \cos(kz - \omega z/v) = \vec{E}_0 \cos(k - n\omega/c)\, z$$

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

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

If the polarization is parallel to optic axis, light is transmitted with one velocity;

if the polarization is perpendicular to optic axis, light is transmitted with different velocity.

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

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

$$\vec{E} = E_x \hat{\varepsilon}_x + E_y \hat{\varepsilon}_y = E_x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

or $$\vec{E}(\vec{r},t) = E_x^0 \cos\left((k - n_x\omega/c)z\right)\hat{\varepsilon}_x + E_y^0 \cos\left((k - n_y\omega/c)z\right)\hat{\varepsilon}_y$$

$$= E^0 \cos 45^\circ \cos\left((k - n_x\omega/c)z\right)\hat{\varepsilon}_x + E^0 \sin 45^\circ \cos\left((k - n_y\omega/c)z\right)\hat{\varepsilon}_y$$

$$= \frac{E^0}{\sqrt{2}}(\cos\left((k - n_x\omega/c)z\right)\hat{\varepsilon}_x + \cos\left((k - n_y\omega/c)z\right)\hat{\varepsilon}_y)$$

This corresponds to representing 45° polarization as superposition of x– and y–polarizations of equal amplitudes, frequency and in phase.

Assume that we choose y–axis to line up with optic axis (-> x–axis perpendicular to optic axis).

This choice is **arbitrary**.

As light passes through, phases change at different rates, i.e., at z' phases will not be equal

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

Thus, if two components start out in phase, **will go in and out of phase** as they travel through the material.

The phase difference is **proportional** to depth into material, i.e.,

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

If thickness is just right to introduce a 90° phase shift between the x– and y–components, linearly polarized (**entering** the material) light will **leave** the material circularly polarized.

Plate with exactly such thickness —> quarter-wave plate. In this case have

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

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

If send light through two such quarter-wave plates, then it exits linearly polarized again in direction at right angles to original direction.

In this case we have

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

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

Example: Birefringence of cellophane

Cellophane consists of long, fibrous molecules.

It is not isotropic since the fibers lie preferentially in one direction.

Create beam of linearly polarized light by sending unpolarized light through sheet of polaroid.

As said earlier, polaroid has useful property that it transmits light that is linearly polarized parallel to preferred axis of polaroid with very little absorption, but light polarized perpendicular to preferred direction is strongly absorbed (not transmitted).

When an unpolarized beam is sent through polaroid, only that part of beam (component of electric field) vibrating parallel to preferred axis of polaroid gets though.

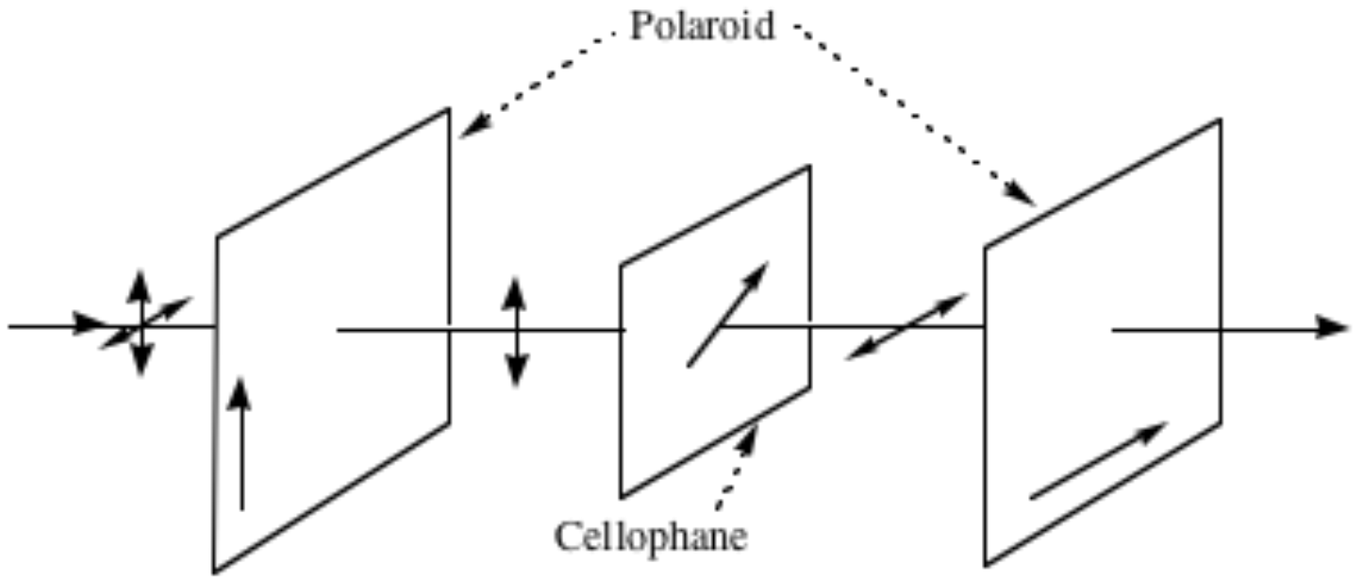
In this manner, the exiting beam is linearly polarized (in the direction of preferred axis).

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

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

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

Birefringence of cellophane illustrated by setup:

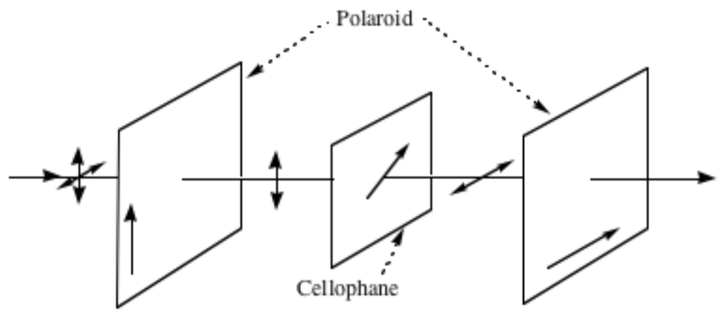


First polaroid produces linearly polarized beam (from entering unpolarized beam).

Linearly polarized beam then passes through cellophane and finally through second polaroid.

Second polaroid detects effect of cellophane on polarization of beam.

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



Now introduce cellophane as shown.

If rotate cellophane sheet around beam axis, find that some light transmits through second polaroid.

In addition, there are two orthogonal orientations of cellophane which permit no light to pass through second polaroid.

These two directions are such that cellophane has no effect on the linear polarization of the beam so that none gets through second polaroid.

The directions are parallel and perpendicular to optic axis of cellophane.

Assume that light passes through cellophane with two different speeds in these two different directions, but is transmitted without changing direction of polarization.

When cellophane is turned halfway between these two directions (as in diagram above) observe that light passing through second polaroid is bright.

Turns out that ordinary cellophane is very close to half-wave thickness for most of colors in white light.

Such sheet will turn the direction of linear polarization through 90° if incident linearly polarized beam makes angle of 45° with the optic axis, so that beam emerging from cellophane is then vibrating in right direction to pass through polaroid sheet.

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

The transmitted color depends on thickness of cellophane. Very neat!

Finally explain polaroids.

Polaroids are materials where not only index of refraction but also amount of absorption is different for light polarized in different directions.

Polaroid consists of thin layer of small crystals of herapathite (salt of iodine and quinine), all aligned with axes parallel (optic axis).

Crystals absorb light when polarization orthogonal to this direction and do not absorb light when polarization is parallel.

If send light into polaroid so that light is linearly polarized at angle θ to optic axis, what intensity will come through?

Simply resolve incident light (electric field) into components parallel($\cos\theta$) and perpendicular($\sin\theta$) to optic axis.

Light that comes out of polaroid is only the cosθ part; the sinθ part is absorbed.

The amplitude that is transmitted is smaller than amplitude that entered.

Entering: $E^0\left(\cos\theta\hat{\varepsilon}_x + \sin\theta\hat{\varepsilon}_y\right)$ Leaving: $E^0\cos\theta\hat{\varepsilon}_x$

Intensity or brightness or energy is:

Entering: $(E^0\cos\theta)^2 + (E^0\sin\theta)^2 = (E^0)^2$ Leaving: $(E^0\cos\theta)^2 = (E^0)^2(\cos\theta)^2 < (E^0)^2$

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

An interesting experiment, which we have already mentioned is the following.

We know that no light will be transmitted through two polaroids if the optic axes are orthogonal.

Now place third polaroid between them with its optic axis at 45° with respect to first polaroid.

Observations show that some light will now be transmitted through second polaroid.

We know that polaroids only absorb light

- they do not create light ,

nevertheless, the addition of the third polaroid allows more light to be transmitted.

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

Only Two Polaroids:

Entering 1st Polaroid: $E^0\left(\cos\theta\hat{\varepsilon}_{x_1}+\sin\theta\hat{\varepsilon}_{y_1}\right)$ Leaving 1st Polaroid: $E^0\cos\theta\hat{\varepsilon}_{x_2}$

Entering 2nd Polaroid: $E^0\cos\theta\hat{\varepsilon}_{x_2}$ Leaving 2nd Polaroid: 0

So no energy (no light) gets through.

Three Polaroids:

Entering 1st Polaroid: $E^0\left(\cos\theta\hat{\varepsilon}_{x_1}+\sin\theta\hat{\varepsilon}_{y_1}\right)$ Leaving 1st Polaroid: $E^0\cos\theta\hat{\varepsilon}_{x_2}$

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

Entering 3rd Polaroid: $E^0\cos\theta\cos 45^\circ\hat{\varepsilon}_{x_2}=\frac{1}{\sqrt{2}}E^0\cos\theta\hat{\varepsilon}_{x_2}$

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

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

Have assumed that x–axis for each polaroid is its optic axis.

All of preceding discussion takes place at classical level.

All phenomena discussed can be explained with classical physics concepts.

As have seen, all such explanations fail at quantum level where ideas like electric fields and components of electric fields will break down - i.e., single photons……

The Quantum Theory of Photon Polarization

Now we carry out the details of a special case that will illustrate how Quantum Mechanics works and also illustrate the mathematical formalism that we have developed earlier.

This discussion is more mathematical than earlier parts of class, but we now have all the tools needed and you will benefit if you persevere and work your way through the material.

As mentioned earlier, the electric field vector $\vec{E}$ of plane electromagnetic waves lies in plane perpendicular to direction of propagation of wave.

Choosing the z–axis as direction of propagation, we can represent electric field vector as a 2–dimensional vector in x – y plane.

This means that we only require two numbers to describe the electric field.

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

$$|\psi\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

where we assume the normalization condition $\langle\psi|\psi\rangle$=1 which —> the condition

$$|\psi_x|^2 + |\psi_y|^2 = 1$$

Examples are:

$$|x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \text{ x - polarized photon(linear or plane polarization)}$$

$$|y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \text{ y - polarized photon(linear or plane polarization)}$$

$$|R\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \rightarrow \text{ right-circular polarized photon}$$

$$|L\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \rightarrow \text{ left-circular polarized photon}$$

$$|45\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \text{ photon linearly polarized at } 45^\circ \text{ to the x-axis}$$

Note that we can write $|R\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\pi/2} \end{pmatrix}$

so relative(difference) phase of components is π/2 as we discussed earlier.

Bra-vector or linear functional <—> ket vector given by row vector

$$\langle\psi| = \begin{pmatrix} \psi_x^* & \psi_y^* \end{pmatrix}$$

which clearly implies via inner product rules

$$\langle\psi \mid \psi\rangle = \begin{pmatrix} \psi_x^* & \psi_y^* \end{pmatrix} \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = |\psi_x|^2 + |\psi_y|^2 = 1$$

In general, for $|\varphi\rangle = \begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix}$ inner product rule says

$$\langle\varphi \mid \psi\rangle = \begin{pmatrix} \varphi_x^* & \varphi_y^* \end{pmatrix} \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \varphi_x^*\psi_x + \varphi_y^*\psi_y = \langle\psi \mid \varphi\rangle^*$$

Also have

$$\langle x \mid x\rangle = 1 = \langle y \mid y\rangle \text{ and } \langle x \mid y\rangle = 0 = \langle y \mid x\rangle \rightarrow \text{ orthonormal set}$$
$$\langle R \mid R\rangle = 1 = \langle L \mid L\rangle \text{ and } \langle R \mid L\rangle = 0 = \langle L \mid R\rangle \rightarrow \text{ orthonormal set}$$

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

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

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

We can find components along the basis vectors using

$$\langle x \mid \psi \rangle = \langle x | \left(\psi_x |x\rangle + \psi_y |y\rangle \right) = \psi_x \langle x \mid x \rangle + \psi_y \langle x \mid y \rangle = \psi_x$$
$$\langle y \mid \psi \rangle = \langle y | \left(\psi_x |x\rangle + \psi_y |y\rangle \right) = \psi_x \langle y \mid x \rangle + \psi_y \langle y \mid y \rangle = \psi_y$$

or

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

and similarly $$|\psi\rangle = |R\rangle \langle R \mid \psi\rangle + |L\rangle \langle L \mid \psi\rangle$$

Basically, we are illustrating examples of the superposition principle

—> any arbitrary polarization state can be written as superposition (linear combination)

of x– and y–polarization states or equivalently,

as superposition of right- and left-circularly polarized states.

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

Classical physics (and now quantum physics) says that beam is superposition of x–polarized beam and y–polarized beam and when beam passes through x–polaroid, effect is to remove y–polarized beam and pass x–polarized beam through unchanged.

Energy of beam given by $|\vec{E}|^2$, which is proportional to $|\psi_x|^2 + |\psi_y|^2$.

Thus, beam energy after passing through x-polaroid proportional to $|\psi_x|^2$.

Fraction of beam energy or fraction of number of photons in beam that passes through given by

$$\frac{|\psi_x|^2}{|\psi_x|^2 + |\psi_y|^2} = |\psi_x|^2 = |\langle x \mid \psi\rangle|^2$$

The earlier discussion for the case of a single photon forced us to set this quantity equal to probability of single photon in state |ψ⟩ passing through x–polaroid or

$$\text{Amplitude that a photon in state } |\psi\rangle \text{ passing through an x-polaroid } = \langle x \,|\, \psi \rangle$$

This agrees with earlier mathematical results.

Using earlier discussions, we also defined ⟨x|ψ⟩ as the probability amplitude for an individual photon to pass through an 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 can write

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

We can generalize polaroid result to say

$$\text{Amplitude that a photon in state } |\psi\rangle$$

$$\text{passes through the prism as a right-circular polarized photon } = \langle R \,|\, \psi\rangle$$

$$\text{Amplitude that a photon in state } |\psi\rangle$$

$$\text{passes through the prism as a left-circular polarized photon } = \langle L \,|\, \psi\rangle$$

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 seen two examples of basis sets for 2–dimensional vector space of polarization states, namely,

$$|x\rangle\,, |y\rangle \quad , \quad |R\rangle\,, |L\rangle$$

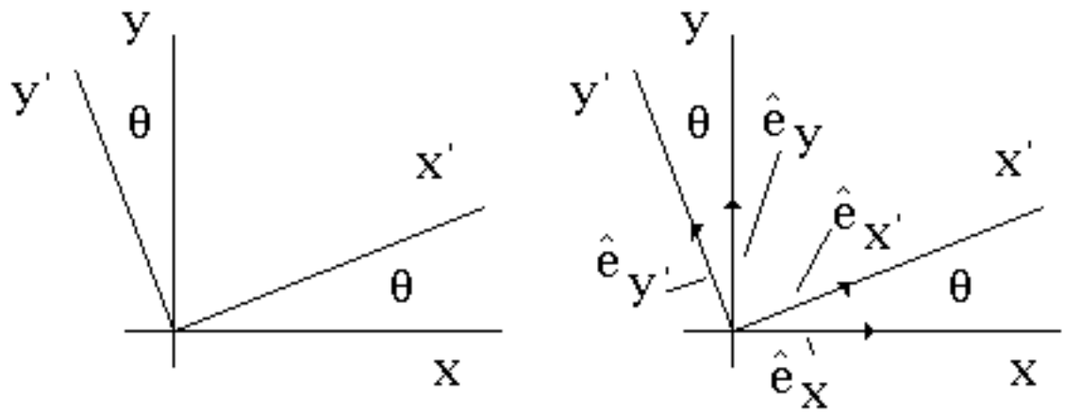
In 2-dimensional vector space there are an infinite number of orthonormal basis sets related to the $|x\rangle$,$|y\rangle$ set.

All are equivalent for describing physical systems (correspond to different orientations of a polaroid in experimental measurements).

Can obtain other sets say, $|x'\rangle$,$|y'\rangle$, by rotation of bases (or axes) as shown in figure (left).

Then have in x–y basis

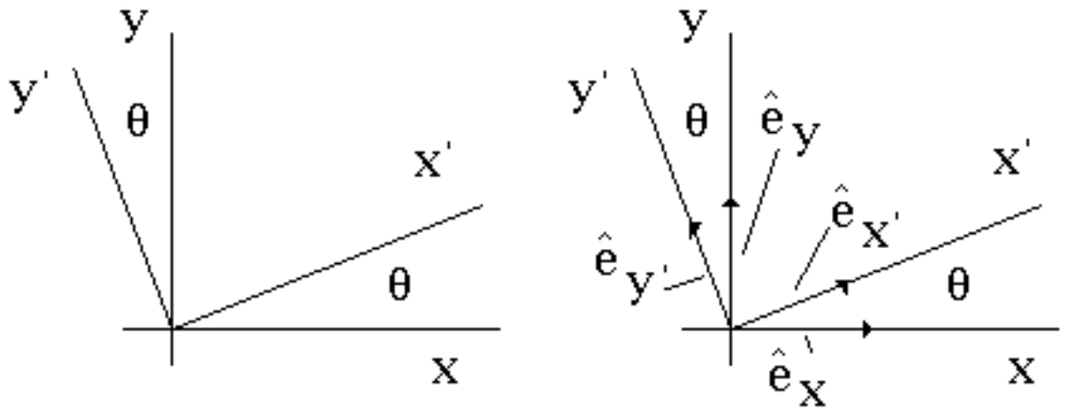
$$|\psi\rangle = \psi_x\,|x\rangle + \psi_y\,|y\rangle = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} = \begin{pmatrix} \langle x \mid \psi\rangle \\ \langle y \mid \psi\rangle \end{pmatrix}$$



and if we choose to use equivalent x′–y′ basis have

$$|\psi\rangle = \psi_{x'}\,|x'\rangle + \psi_{y'}\,|y'\rangle = \begin{pmatrix} \psi_{x'} \\ \psi_{y'} \end{pmatrix} = \begin{pmatrix} \langle x' \mid \psi\rangle \\ \langle y' \mid \psi\rangle \end{pmatrix}$$

How are these components related to each other?



Have from earlier that

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

which implies

$$\langle x' \mid \psi\rangle = \langle x' \mid x\rangle \langle x \mid \psi\rangle + \langle x' \mid y\rangle \langle y \mid \psi\rangle$$
$$\langle y' \mid \psi\rangle = \langle y' \mid x\rangle \langle x \mid \psi\rangle + \langle y' \mid y\rangle \langle y \mid \psi\rangle$$

or using matrix multiplication notation

$$\begin{pmatrix} \langle x' \mid \psi\rangle \\ \langle y' \mid \psi\rangle \end{pmatrix} = \begin{pmatrix} \langle x' \mid x\rangle & \langle x' \mid y\rangle \\ \langle y' \mid x\rangle & \langle y' \mid y\rangle \end{pmatrix} \begin{pmatrix} \langle x \mid \psi\rangle \\ \langle y \mid \psi\rangle \end{pmatrix}$$

So we can transform basis (transform components) if we can determine the 2 × 2 transformation matrix

$$\begin{pmatrix} \langle x' \mid x\rangle & \langle x' \mid y\rangle \\ \langle y' \mid x\rangle & \langle y' \mid y\rangle \end{pmatrix}$$

that is used in the above equation.

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

For linear(plane) polarized case, we can think of an analogy to unit vectors along axes in ordinary space as shown in above figure on right.

Then have(by analogy)

$$\hat{e}_x \cdot \hat{e}_{x'} = \cos\theta = \langle x' \mid x \rangle \quad , \quad \hat{e}_{x'} \cdot \hat{e}_y = \sin\theta = \langle x' \mid y \rangle$$
$$\hat{e}_x \cdot \hat{e}_{y'} = \cos\theta = \langle y' \mid y \rangle \quad , \quad \hat{e}_{y'} \cdot \hat{e}_x = -\sin\theta = \langle y' \mid x \rangle$$

or

$$|x\rangle = \langle x \mid x' \rangle |x'\rangle + \langle x \mid y' \rangle |y'\rangle = \cos\theta |x'\rangle - \sin\theta |y'\rangle$$
$$|y\rangle = \langle y \mid x' \rangle |x'\rangle + \langle y \mid y' \rangle |y'\rangle = \sin\theta |x'\rangle + \cos\theta |y'\rangle$$

or

$$\begin{pmatrix} \langle x' \mid \psi \rangle \\ \langle y' \mid \psi \rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \langle x \mid \psi \rangle \\ \langle y \mid \psi \rangle \end{pmatrix} = \hat{R}(\theta) \begin{pmatrix} \langle x \mid \psi \rangle \\ \langle y \mid \psi \rangle \end{pmatrix}$$

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

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

There are two equivalent ways to interpret these results.

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

Second, can rotate vector and keep axes fixed(rotate in opposite direction).

In this case, we regard

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

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

$$\langle x' \mid \psi \rangle = \langle x \mid \psi' \rangle \, , \langle y' \mid \psi \rangle = \langle y \mid \psi' \rangle$$

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

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

It is transformation of vectors and is unitary operator. Can see this as follows:

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

Transformation operators are unitary because transform state vectors which must not change lengths (otherwise probability ideas get messed up).

This follows from fact that unitary transformations preserve inner products(and so they preserve lengths)

$$|\psi'\rangle = \hat{R}(\theta) \, |\psi\rangle \quad , \quad |\varphi'\rangle = \hat{R}(\theta) \, |\varphi\rangle$$
$$\langle \varphi' \mid \psi' \rangle = \langle \varphi | \, \hat{R}^+(\theta) \hat{R}(\theta) \, |\psi\rangle = \langle \varphi | \, \hat{R}^{-1}(\theta) \hat{R}(\theta) \, |\psi\rangle = \langle \varphi | \, \hat{I} \, |\psi\rangle = \langle \varphi \mid \psi \rangle$$

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

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

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

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

where $$\hat{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{ Identity operator} \quad , \quad \hat{Q} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

and physical meaning of operator $\hat{Q}$ is yet to be determined

Now show that $\hbar\hat{Q} = \hat{J}_z$.

Expanding exponential in power series have

$$\hat{R}(\theta) = e^{i\theta \frac{\hat{J}_z}{\hbar}} = \hat{I} + (i\frac{\hat{J}_z}{\hbar})\theta + \frac{(i\frac{\hat{J}_z}{\hbar})^2}{2!}\theta^2 + \frac{(i\frac{\hat{J}_z}{\hbar})^3}{3!}\theta^3 + \frac{(i\frac{\hat{J}_z}{\hbar})^4}{4!}\theta^4 + ...$$

$$= \hat{R}(0) + \frac{1}{1!}\left.\frac{d\hat{R}(\theta)}{d\theta}\right|_{\theta=0}\theta + \frac{1}{2!}\left.\frac{d^2\hat{R}(\theta)}{d\theta^2}\right|_{\theta=0}\theta^2 + \frac{1}{3!}\left.\frac{d^3\hat{R}(\theta)}{d\theta^3}\right|_{\theta=0}\theta^3 + \frac{1}{4!}\left.\frac{d^4\hat{R}(\theta)}{d\theta^4}\right|_{\theta=0}\theta^4 + ...$$

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

$$\hat{R}(\theta) = e^{i\theta\frac{\hat{J}_z}{\hbar}} = \hat{I} + (i\frac{\hat{J}_z}{\hbar})\theta + \frac{(i\frac{\hat{J}_z}{\hbar})^2}{2!}\theta^2 + \frac{(i\frac{\hat{J}_z}{\hbar})^3}{3!}\theta^3 + \frac{(i\frac{\hat{J}_z}{\hbar})^4}{4!}\theta^4 + ...$$

$$= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - .....\right)\hat{I} + i\left(\frac{\theta}{1!} - \frac{\theta^3}{3!} + ...\right)\hat{Q} \quad = \cos\theta\hat{I} + i\sin\theta\hat{Q}$$

which agrees with earlier result. Thus, have

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

must be matrix representing angular momentum operator $\hat{J}_z$ in |x⟩ , |y⟩ basis. This completes the proof.

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

$$\hat{R}(\theta)\,|\psi\rangle = (\cos\theta\hat{I} + i\sin\theta\hat{Q})\,|\psi\rangle$$

$$\hat{R}(\theta)\,|\psi\rangle = (\cos\theta\hat{I} + i\sin\theta\frac{\hat{J}_z}{\hbar})\,|\psi\rangle = c\,|\psi\rangle$$

where c = eigenvalue corresponding to eigenvector |ψ⟩.

Since all vectors are eigenvectors of identity operator $\hat{I}$, only need to find eigenvectors and eigenvalues of $\hat{J}_z$ in order to solve problem for $\hat{R}(\theta)$ (have same eigenvectors since only differ by identity operator).

Let

$$\hat{J}_z \left|\psi\right\rangle = \lambda \left|\psi\right\rangle$$

Now, since $\hat{J}_z^2 = \hbar^2 \hat{I}$, i.e.,

$$\hat{J}_z^2 = \hbar^2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \hbar^2 \hat{I}$$

we have

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

which says that

$$\lambda^2 = \hbar^2 \quad \rightarrow \quad \lambda = \pm\hbar = eigenvalues\ of\ \hat{J}_z$$

We can find the corresponding eigenvectors by inserting eigenvalues into eigenvalue/ eigenvector equation $\hat{J}_z \left|J_z = \hbar\right\rangle = \hbar \left|J_z = \hbar\right\rangle$

Assume that

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

then we get

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

—> result ia = b, which together with normalization condition says that $a = 1/\sqrt{2}$

Have arbitrarily chosen a to be real since only relative phase between components will be important in quantum mechanics.

This then gives $b = i/\sqrt{2}$.

Finally have eigenvector

$$|J_z = \hbar\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = |R\rangle \qquad \text{Similarly, get} \qquad |J_z = -\hbar\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = |L\rangle$$

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

$$\hat{R}(\theta)\,|R\rangle = (\cos\theta \hat{I} + i \sin\theta \tfrac{\hat{J}_z}{\hbar})\,|R\rangle = (\cos\theta + i\sin\theta)\,|R\rangle = e^{i\theta}\,|R\rangle$$
$$\hat{R}(\theta)\,|L\rangle = e^{-i\theta}\,|L\rangle$$

Physically, —> |R⟩ and |L⟩ states are only changed by overall phase factor under rotation of the basis.

This allows us to specify what happens to arbitrary vector |ψ⟩ under rotations.

First, expand arbitrary vector in |R⟩,|L⟩ basis. Then apply rotation operator to obtain

$$\hat{R}(\theta)\,|\psi\rangle = \hat{R}(\theta)\,|R\rangle\,\langle R \mid \psi\rangle + \hat{R}(\theta)\,|L\rangle\,\langle L \mid \psi\rangle = e^{i\theta}\,|R\rangle\,\langle R \mid \psi\rangle + e^{-i\theta}\,|L\rangle\,\langle L \mid \psi\rangle$$

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

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

Now, it is experimental fact that if a photon traveling in 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 same, nor does it change by any value other than $\pm\hbar$

We interpret these results to say that RCP photon is in state which is eigenvector of $\hat{J}_z$ with eigenvalue $\hbar$ for that photon in state has spin = $\hbar$.

Similarly, LCP photon has spin = $-\hbar$.

One cannot predict, for any single photon, whether change will be $+\hbar$ or $-\hbar$.

Can, however, predict probability of either value occurring.

In particular, according to probability formalism, must have

$$|\langle R \mid \psi \rangle|^2 = \text{ probability of } +\hbar$$
$$|\langle L \mid \psi \rangle|^2 = \text{ probability of } -\hbar$$

and average value of z–component of angular momentum is

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

$$\left\langle \hat{J}_z \right\rangle = \hbar |\langle R \mid \psi \rangle|^2 - \hbar |\langle L \mid \psi \rangle|^2$$

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

We can still talk in terms of probabilities, however.

The discreteness of angular momentum spectrum once again forces probabilistic interpretation on us.

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

$$\left\langle \hat{J}_z \right\rangle = \langle\psi| \hat{J}_z |\psi\rangle$$

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

$$\left\langle \hat{J}_z \right\rangle = (\langle R \mid \psi\rangle^* \langle R| + \langle L \mid \psi\rangle^* \langle L|)\hat{J}_z(|R\rangle \langle R \mid \psi\rangle + |L\rangle \langle L \mid \psi\rangle)$$

$$= \langle R| \hat{J}_z |R\rangle |\langle R \mid \psi\rangle|^2 + \langle L| \hat{J}_z |L\rangle |\langle L \mid \psi\rangle|^2$$

$$+ \langle R| \hat{J}_z |L\rangle \langle R \mid \psi\rangle^* \langle L \mid \psi\rangle + \langle L \mid \psi\rangle^* \langle R \mid \psi\rangle \langle L| \hat{J}_z |R\rangle$$

$$= \hbar |\langle R \mid \psi\rangle|^2 - \hbar |\langle L \mid \psi\rangle|^2$$ as above

Let us return for moment to matrix representation of operator. Have found following results:

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

Thus, in {|R⟩,|L⟩} basis, these relations imply 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 standard form of $\hat{J}_z$ in terms of one of so-called Pauli matrices, namely

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

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

and, therefore, in {|x⟩,|y⟩} basis, have matrix representation

$$\hat{J}_z = \begin{pmatrix} \langle x| \hat{J}_z |x\rangle & \langle x| \hat{J}_z |y\rangle \\ \langle y| \hat{J}_z |x\rangle & \langle y| \hat{J}_z |y\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$ —> same form as derived earlier.

Projection Operators

Now let us look at projection operators in the context of photon polarization.

The projection operator $|\psi\rangle\langle\phi|$ can be represented by a 2 × 2 matrix in the polarization state vector space.

It is constructed using a mathematical object called an **outer product**.

$$\hat{P} = |\psi\rangle\langle\varphi| = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} \begin{pmatrix} \varphi_x^* & \varphi_y^* \end{pmatrix} = \begin{pmatrix} \psi_x\varphi_x^* & \psi_x\varphi_y^* \\ \psi_y\varphi_x^* & \psi_y\varphi_y^* \end{pmatrix}$$

or equivalently, by choosing basis and finding matrix representation

$$\hat{P} = \begin{pmatrix} \langle x|\hat{P}|x\rangle & \langle x|\hat{P}|y\rangle \\ \langle y|\hat{P}|x\rangle & \langle y|\hat{P}|y\rangle \end{pmatrix} = \begin{pmatrix} \langle x\mid\psi\rangle\langle\varphi\mid x\rangle & \langle x\mid\psi\rangle\langle\varphi\mid y\rangle \\ \langle y\mid\psi\rangle\langle\varphi\mid x\rangle & \langle y\mid\psi\rangle\langle\varphi\mid y\rangle \end{pmatrix} = \begin{pmatrix} \psi_x\varphi_x^* & \psi_x\varphi_y^* \\ \psi_y\varphi_x^* & \psi_y\varphi_y^* \end{pmatrix}$$

In particular, we have in the {|x⟩,|y⟩} basis

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

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

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

as specified in earlier mathematical discussions.

Similarly, we have $\hat{I} = |R\rangle\langle R| + |L\rangle\langle L|$

which leads to $\hat{J}_z = \hat{J}_z\hat{I} = \hat{J}_z|R\rangle\langle R| + \hat{J}_z|L\rangle\langle L| = \hbar|R\rangle\langle R| - \hbar|L\rangle\langle L|$

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

The action of the polarizer can be considered as a measurement.

What are operators representing such measurements?

Clearly, operators for x– and y–polarizers are given by

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

since

$$\hat{O}_x|\psi\rangle = \hat{O}_x\left(a|x\rangle + b|y\rangle\right) = \left(|x\rangle\langle x|\right)\left(a|x\rangle + b|y\rangle\right) = a|x\rangle \qquad \textbf{picks out x part!}$$

and so on.

If light polarized at angle θ from x–axis, it is in state $|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$

Operator representing polarizer at angle θ is (in 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 probability of measuring x–polarization when in θ state is

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

A result we have seen several times in earlier discussions.

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 linear combination of basis states

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

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

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

Might now conclude, from experimental fact that only $\pm\hbar$ is transferred to matter, that photons always either in state $|R\rangle$ with probability α or in state $|L\rangle$ with probability 1 – α.

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

FACT: 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 probability = $|\langle R \mid x\rangle|^2$ of being RCP and RCP photon has probability = $|\langle y \mid R\rangle|^2$ of being y–polarized photon and thus passing through y–polaroid.

(b) an x–polarized photon has probability = $|\langle L \mid x\rangle|^2$ of being LCP and LCP photon has probability = $|\langle y \mid L\rangle|^2$ of being y–polarized photon and thus passing through y–polaroid.

This means that total probability that x–polarized photon would get through y–polaroid in this interpretation is

$$\text{total probability } = |\langle R \mid x\rangle|^2 \, |\langle y \mid R\rangle|^2 + |\langle L \mid x\rangle|^2 \, |\langle y \mid L\rangle|^2 = \frac{1}{2}$$

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

SOLUTION: When we think of an x–polarized photon being RCP photon or 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 word interference here in this way.

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

(a) Probability amplitude of x–polarized photon passing through y–polaroid = $\langle y \mid x\rangle = 0$, which implies that probability = $|\langle y \mid x\rangle|^2 = 0$ also.

(b) If we say that an x–polarized photon is in a superposition of $|R\rangle$ and $|L\rangle$ (make no statement about probabilities at this point), this implies that

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

which gives $$\langle y \mid x\rangle = \langle y \mid R\rangle \langle R \mid x\rangle + \langle y \mid L\rangle \langle L \mid x\rangle$$

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 an RCP photon $\langle y|R\rangle\langle R|x\rangle$ and that passes through as an LCP photon $\langle y|L\rangle\langle L|x\rangle$. **This does not say that it has actually done either!**

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

$$
\begin{aligned}
probability &= |\langle y \mid R\rangle \langle R \mid x\rangle + \langle y \mid L\rangle \langle L \mid x\rangle|^2 \\
&= \left(\langle y \mid R\rangle^* \langle R \mid x\rangle^* + \langle y \mid L\rangle^* \langle L \mid x\rangle^*\right)\left(\langle y \mid R\rangle \langle R \mid x\rangle + \langle y \mid L\rangle \langle L \mid x\rangle\right) \\
&= |\langle y \mid R\rangle|^2 |\langle R \mid x\rangle|^2 + |\langle y \mid L\rangle|^2 |\langle L \mid x\rangle|^2 \\
&\qquad + \langle y \mid R\rangle \langle R \mid x\rangle \langle y \mid L\rangle^* \langle L \mid x\rangle^* + \langle y \mid R\rangle^* \langle R \mid x\rangle^* \langle y \mid L\rangle \langle L \mid x\rangle
\end{aligned}
$$

(d) The first two terms are same as incorrect calculation done earlier.

The last two terms represent interference effects between the two amplitudes (RCP way and LCP way).

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

INTERPRETATION: The way to interpret this result is as follows:

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

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

These are indistinguishable ways for process to occur, i.e., no measurement exists that can tell us whether it passes through as an RCP photon or as an LCP photon without destroying the interference, i.e., without radically altering the experiment.

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

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

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

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

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

$$\langle y \mid R \rangle \langle R \mid x \rangle$$

(2) Total amplitude for process that can take place in several indistinguishable ways is sum of amplitudes for each individual way, i.e.,

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

Note here that this is merely a reflection of property of projection operators that

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

which says that

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

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

(3) Total probability for process to occur is absolute square of total amplitude.

So, in classical physics, we

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

——> 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

——> get interference effects!!

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 two ways actually happened without altering experiment.

(2) In particular, if we attempt to find out, then the interference effects will disappear and we will return to classical result obtained by adding probabilities.

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

This happens as follows: instead of

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

if we attempted to add measurement to determine if x–polarized photon was RCP or LCP, we would have

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

Probability calculation then give

$$total\ probability = |\langle y \mid R\rangle|^2 |\langle R \mid x\rangle|^2 + |\langle y \mid L\rangle|^2 |\langle L \mid x\rangle|^2$$

$$+2\mathrm{Real}\left[\langle y \mid R\rangle \langle R \mid x\rangle e^{i(\alpha_R - \alpha_L)} \langle y \mid L\rangle^* \langle L \mid x\rangle^*\right]$$

Observed probability, which is result of many identical measurements in laboratory, is average over all values of extra phases(they are random).

This involves integrating over 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 interference term averages to zero and we get the classical result!!

More about Pure States, Unpure States and Density Operators

If photon were in state |x⟩, then would have, for some linear operator $\hat{A}$, an expectation value or average value given by

$$\left\langle \hat{A} \right\rangle = \langle x| \hat{A} |x\rangle$$

We defined a property of an operator called trace as

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

that is, sum over diagonal matrix elements.

Some Properties of the Trace:

$$Tr(\hat{A}\hat{B}) = Tr(\hat{B}\hat{A})$$
$$Tr(c\hat{B}) = cTr(\hat{B})$$
$$Tr(c(\hat{A} + \hat{B})) = Tr(c\hat{A}) + Tr(c\hat{B}) = cTr(\hat{A}) + cTr(\hat{B})$$

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

$$Tr\hat{W} = 1 = \sum_k W_{kk} = \sum_k \langle \varphi_k | \hat{W} |\varphi_k\rangle$$

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

Quantum theory assumes(equivalent postulates)

(1) A density operator exists for every real physical system (in same way that every physical system can be represented by state vector or ket).

(2) The expectation value of an operator $\hat{A}$ is given by $\left\langle \hat{A} \right\rangle = Tr(\hat{W}\hat{A})$

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

In particular, choose as density operator $\hat{W}$ projection operator for some vector (—> pure state)

$$\hat{W} = |\psi\rangle\langle\psi|$$

$\hat{W}^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi \mid \psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{W}$ —> idempotent operator

and thus has eigenvalues $w_k = 0, 1$ only, i.e.,

$$\hat{W}|\beta\rangle = \beta|\beta\rangle \rightarrow \hat{W}^2|\beta\rangle = \hat{W}\beta|\beta\rangle = \beta^2|\beta\rangle = \hat{W}|\beta\rangle = \beta|\beta\rangle$$
$$\rightarrow \left(\beta^2 - \beta\right)|\beta\rangle = 0 \rightarrow \beta^2 - \beta = 0 \rightarrow \beta = 0, 1$$

Assume that eigenvector corresponding to eigenvalue 1 is $|\alpha\rangle$.

Properties of the Density Operator

$$\sum_k w_k = 0 + 1 = 1 = Tr\hat{W}$$
$$\langle a|\hat{W}|a\rangle = |\langle a \mid \alpha\rangle|^2 \geq 0$$

so that all required properties for density operator are, in fact, satisfied by the assumed form.

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

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

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

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

Since $\hat{W}$ is Hermitian, its eigenvectors must form an orthonormal basis where $\langle w_k \mid w_j \rangle = \delta_{kj}$

Now derive some other properties of this density operator object.

Spectrum of $\hat{W}$ is discrete set of numbers {w_k}. Then have

$$Tr\hat{W} = 1 = \sum_j \langle w_j | \hat{W} | w_j \rangle = \sum_j \langle w_j | w_j | w_j \rangle = \sum_j w_j \langle w_j \mid w_j \rangle$$
$$\rightarrow \sum_j w_j = 1$$

Since $\hat{W}$ is Hermitian, have $\hat{W} = \hat{W}^+$ —> eigenvalues are real numbers $w_k = w_k^*$

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

$$\langle a | \hat{W} | a \rangle = \langle a | \sum_k w_k | w_k \rangle \langle w_k | a \rangle = \sum_k w_k \langle a \mid w_k \rangle \langle w_k \mid a \rangle = \sum_k w_k |\langle a \mid w_k \rangle|^2 \geq 0$$

for any vector |a⟩.

Can only be true, in general, if $w_k \geq 0$ for all k.

Results $w_k \geq 0$, $\sum_k w_k = 1$ imply that $0 \leq w_k \leq 1$

Returning to simple case of pure state $\hat{W} = |\psi\rangle \langle\psi|$, then have

$$\langle \hat{B} \rangle = \langle \psi | \hat{B} | \psi \rangle = \langle \psi | \hat{B} \hat{I} | \psi \rangle$$
$$= \langle \psi | \hat{B} \left(\sum_k |w_k\rangle \langle w_k| \right) |\psi\rangle = \sum_k \langle \psi | \hat{B} | w_k \rangle \langle w_k \mid \psi \rangle$$
$$= \sum_k \langle w_k \mid \psi \rangle \langle \psi | \hat{B} | w_k \rangle = \sum_k \langle w_k | \left(|\psi\rangle \langle \psi | \hat{B} | w_k \right\rangle = \sum_k \langle w_k | \hat{W} \hat{B} | w_k \rangle = Tr(\hat{W}\hat{B})$$

Since important quantities for connection to experiment will be these expectation values, we see that state represented by $\hat{W}$ is equally well represented by state vector $|\psi\rangle$ in this 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 state is pure or not follows from 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 $|x\rangle$, $|y\rangle$ basis have

$$\left\langle \hat{A} \right\rangle = \langle x| \hat{A} |x\rangle = Tr(\hat{W}\hat{A}) = \langle x| \hat{W}\hat{A} |x\rangle + \langle y| \hat{W}\hat{A} |y\rangle$$

$$= \langle x| \hat{W}\hat{I}\hat{A} |x\rangle + \langle y| \hat{W}\hat{I}\hat{A} |y\rangle$$

$$= \langle x| \hat{W} |x\rangle \langle x| \hat{A} |x\rangle + \langle x| \hat{W} |y\rangle \langle y| \hat{A} |x\rangle + \langle y| \hat{W} |x\rangle \langle x| \hat{A} |y\rangle + \langle y| \hat{W} |y\rangle \langle y| \hat{A} |y\rangle$$

This implies that $\langle x|\,\hat{W}\,|x\rangle = 1 \quad , \quad \langle x|\,\hat{W}\,|y\rangle = \langle y|\,\hat{W}\,|x\rangle = \langle y|\,\hat{W}\,|y\rangle = 0$ or

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

which is linear combination(one item in this case) of density operators and —> a pure state.

Now suppose that photon is in state

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

where know that phases are equal, $\alpha_x = \alpha_y$ (relative phase between components known exactly in state) so that

$$|\psi\rangle = e^{i\alpha_x}\left(\frac{1}{\sqrt{2}}|x\rangle + \frac{1}{\sqrt{2}}|y\rangle\right)$$ Terms $e^{i\alpha_x}$, etc, called phase factors.

But, since all states must have length 1, can ignore overall phase factor and write

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

This says that probability = 1/2 that photon behaves like |x⟩ and probability = 1/2 that photon behaves like |y⟩. In this case, have

$$\left\langle \hat{A} \right\rangle = \langle\psi|\,\hat{A}\,|\psi\rangle = \frac{1}{2}\left[\langle x|\,\hat{A}\,|x\rangle + \langle x|\,\hat{A}\,|y\rangle + \langle y|\,\hat{A}\,|x\rangle + \langle y|\,\hat{A}\,|y\rangle\right]$$

$$= Tr(\hat{W}\hat{A}) = \langle x|\,\hat{W}\hat{A}\,|x\rangle + \langle y|\,\hat{W}\hat{A}\,|y\rangle = \langle x|\,\hat{W}\hat{I}\hat{A}\,|x\rangle + \langle y|\,\hat{W}\hat{I}\hat{A}\,|y\rangle$$

$$= \langle x|\,\hat{W}\,|x\rangle\langle x|\,\hat{A}\,|x\rangle + \langle x|\,\hat{W}\,|y\rangle\langle y|\,\hat{A}\,|x\rangle + \langle y|\,\hat{W}\,|x\rangle\langle x|\,\hat{A}\,|y\rangle + \langle y|\,\hat{W}\,|y\rangle\langle y|\,\hat{A}\,|y\rangle$$

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

So, again have pure state.

But what happens **if only know** that probability = 1/2 that photon behaves like |x⟩ and probability = 1/2 that photon behaves like |y⟩.

—> state vector is $$|\psi\rangle = a\,|x\rangle + b\,|y\rangle$$

where only know that $|a|^2 = |b|^2 = 1/2$. Let us choose $$a = \frac{e^{i\alpha_a}}{\sqrt{2}} \quad , \quad b = \frac{e^{i\alpha_b}}{\sqrt{2}}$$

Do not have any phase information in this case.

In addition, phase values could be different in each separate experiment

—> must average over relative phase αx – αy when computing probabilities

and thus all interference effects will vanish as shown below.

When calculate expectation value have

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

and when average over relative phase obtain $$\left\langle \hat{A} \right\rangle = \frac{1}{2}\,\langle x|\,\hat{A}\,|x\rangle + \frac{1}{2}\,\langle y|\,\hat{A}\,|y\rangle$$

Again, must have

$$\left\langle \hat{A} \right\rangle = Tr(\hat{W}\hat{A}) = \langle x|\,\hat{W}\hat{A}\,|x\rangle + \langle y|\,\hat{W}\hat{A}\,|y\rangle = \langle x|\,\hat{W}\hat{I}\hat{A}\,|x\rangle + \langle y|\,\hat{W}\hat{I}\hat{A}\,|y\rangle$$

$$= \langle x|\,\hat{W}\,|x\rangle\,\langle x|\,\hat{A}\,|x\rangle + \langle x|\,\hat{W}\,|y\rangle\,\langle y|\,\hat{A}\,|x\rangle + \langle y|\,\hat{W}\,|x\rangle\,\langle x|\,\hat{A}\,|y\rangle + \langle y|\,\hat{W}\,|y\rangle\,\langle y|\,\hat{A}\,|y\rangle$$

which implies that

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

or

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

This is a nonpure state.

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

Unpolarized Light

Consider following experiment:

Have beam of monochromatic light composed of photons from two sources which output photons in states $|\psi_1\rangle$ or $|\psi_2\rangle$, respectively.

Sources emit photons randomly and they are independent of each other —> cannot tell which source a particular photon comes from.

Assign these probabilities

p_1 = probability that photon comes from source #1

p_2 = probability that photon comes from source #2

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

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

and probability that it transfers $-\hbar$ is

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

—> average value of angular momentum transfer for beam of photons is

$$\left\langle \hat{J}_z \right\rangle = \hbar p_+ - \hbar p_- = \hbar p_1 \left|\langle R \mid \psi_1 \rangle\right|^2 + \hbar p_2 \left|\langle R \mid \psi_2 \rangle\right|^2 - \hbar p_1 \left|\langle L \mid \psi_1 \rangle\right|^2 - \hbar p_2 \left|\langle L \mid \psi_2 \rangle\right|^2$$

$$= p_1 \left[\hbar \left|\langle R \mid \psi_1 \rangle\right|^2 - \hbar \left|\langle L \mid \psi_1 \rangle\right|^2\right] + p_2 \left[\hbar \left|\langle R \mid \psi_2 \rangle\right|^2 - \hbar \left|\langle L \mid \psi_2 \rangle\right|^2\right]$$

$$= p_1 \left\langle \hat{J}_z \right\rangle_1 + p_2 \left\langle \hat{J}_z \right\rangle_2$$

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

Emphasize that it is important to realize that statement photon is either in state but do not know which is NOT same statement as photon is in state which is superposition of $|\psi_1\rangle$ and $|\psi_2\rangle$.

In second case, saying relative phase is known as in state $|\psi\rangle = \frac{1}{\sqrt{2}} |x\rangle + \frac{1}{\sqrt{2}} |y\rangle$

which was found to be a pure state.

Being in superposition implies that know relative phase of components.

In first case, however, saying that relative phase unknown and —> interference effects will vanish.

In pure states, have superpositions and probability amplitude rules apply.

In nonpure or mixed states, where system is in one of several states with definite probabilities, find weighted averages (weighted with state probabilities) of value in each state.

Use addition of probabilities with no interference effects —> equivalent to saying 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 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.

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

In particular, consider example of birefringent crystal, such as calcite.

Calcite crystal has preferred direction called 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.

Assume that optic axis is in x-y plane and send beam of photons in z-direction.

Photons polarized perpendicular to optic axis called **ordinary** and are in state $|o\rangle$ and photons polarized parallel to optic axis called **extraordinary** and are in state $|e\rangle$.

Set of states $\{|o\rangle,|e\rangle\}$ forms an orthonormal basis and general photon states interacting with calcite crystal are written as superpositions of these basis states —> example of a general rule in quantum mechanics.

If doing experiment using particular measuring device that measures observable $\hat{Q}$, then should use as basis for all states, eigenvectors of $\hat{Q}$.

This requirement pushes us to ask correct experimental questions (those that quantum mechanics can answer).

This particular basis is called the home space for experiment.

Now, as we saw earlier, the phase of light wave with wavelength λ as propagates through medium in z–direction given by quantity $\phi = e^{ikz}$ with

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

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 calcite crystal is to change relative phase of $|o\rangle$ and $|e\rangle$ components making up superposition.

Assume that state of photon entering the calcite crystal is

$$|\psi_{in}\rangle = |e\rangle\langle e \mid \psi_{in}\rangle + |o\rangle\langle o \mid \psi_{in}\rangle$$

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

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

$$|\psi_{out}\rangle = e^{ik_e\ell}\,|e\rangle\,\langle e \mid \psi_{in}\rangle + e^{ik_o\ell}\,|o\rangle\,\langle o \mid \psi_{in}\rangle = \hat{U}_\ell\,|\psi_{in}\rangle$$

where

$$\hat{U}_z = e^{ik_e z}\,|e\rangle\,\langle e| + e^{ik_o z}\,|o\rangle\,\langle o|$$

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

Now define two new quantities which are very important throughout the study of Quantum Mechanics.

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

(1) **transition amplitude** for photon to enter calcite in state $|\psi_{in}\rangle$ and leave in state $|\phi\rangle$ is

$$\langle\phi \mid \psi_{out}\rangle = \langle\phi|\,\hat{U}_\ell\,|\psi_{in}\rangle$$

(2) **transition probability** is $$|\langle\phi \mid \psi_{out}\rangle|^2 = \left|\langle\phi|\,\hat{U}_\ell\,|\psi_{in}\rangle\right|^2$$

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

$$|\psi_z\rangle = \text{ state of the photon after traveling distance } z \text{ in calcite } = \hat{U}_z\,|\psi_{in}\rangle$$

From form of $\hat{U}_z$ have

$$\hat{U}_{z+\varepsilon} = e^{ik_e(z+\varepsilon)}\,|e\rangle\,\langle e| + e^{ik_o(z+\varepsilon)}\,|o\rangle\,\langle o| \;= (e^{ik_e\varepsilon}\,|e\rangle\,\langle e| + e^{ik_o\varepsilon}\,|o\rangle\,\langle o|)(e^{ik_e z}\,|e\rangle\,\langle e| + e^{ik_o z}\,|o\rangle\,\langle o|)$$

or $\hat{U}_{z+\varepsilon} = \hat{U}_\varepsilon \hat{U}_z$ —> general result for time development operators, namely $\hat{U}_{t+t'} = \hat{U}_{t'} \hat{U}_t$

—-> $|\psi_{z+\varepsilon}\rangle = \hat{U}_{z+\varepsilon} |\psi_{in}\rangle = \hat{U}_\varepsilon \hat{U}_z |\psi_{in}\rangle = \hat{U}_\varepsilon |\psi_z\rangle$

Now let $\varepsilon \to 0$ such that $k_0\varepsilon \ll 1$ and $k_e\varepsilon \ll 1$ and we can write (to 1st order)

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

where $\hat{I} = |e\rangle \langle e| + |o\rangle \langle o| \quad , \quad \hat{K} = k_e |e\rangle \langle e| + k_o |o\rangle \langle o|$

Now, the relation

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

is an expansion of an operator in terms of its eigenvalues and corresponding projection operators (eigenvectors).

It says that eigenvectors of $\hat{K}$ are $|e\rangle$ and $|o\rangle$ with eigenvalues k_e and k_o, respectively.

This illustrates awesome power in these methods!!

Then we have

$$|\psi_{z+\varepsilon}\rangle = \left(\hat{I} + i\varepsilon\hat{K}\right) |\psi_z\rangle \quad \text{or} \quad |\psi_{z+\varepsilon}\rangle - |\psi_z\rangle = i\varepsilon\hat{K} |\psi_z\rangle$$

or $\lim_{\varepsilon \to 0} \frac{|\psi_{z+\varepsilon}\rangle - |\psi_z\rangle}{\varepsilon} = i\hat{K} |\psi_z\rangle$

—> differential equation for time development of state vector $\frac{d}{dz} |\psi_z\rangle = i\hat{K} |\psi_z\rangle$

Similar to differential equation obtained earlier for time development operator —> have

$$\hat{K} = \text{ hermitian operator} \quad , \quad \hat{U}_z = \text{ unitary operator}$$

Derive some important results. We have, using the x – y basis

$$\langle x \mid \psi_{z+\varepsilon} \rangle - \langle x \mid \psi_z \rangle = i\varepsilon \langle x| \hat{K} |\psi_z\rangle = i\varepsilon \langle x| \hat{K}\hat{I} |\psi_z\rangle$$

$$= i\varepsilon \langle x| \hat{K} |x\rangle \langle x \mid \psi_z\rangle + i\varepsilon \langle x| \hat{K} |y\rangle \langle y \mid \psi_z\rangle$$

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

Similarly, have

$$\langle y \mid \psi_{z+\varepsilon} \rangle - \langle y \mid \psi_z \rangle = i\varepsilon \langle y| K |\psi_z\rangle = i\varepsilon \langle y| KI |\psi_z\rangle$$

$$= i\varepsilon \langle y| \hat{K} |x\rangle \langle x \mid \psi_z\rangle + i\varepsilon \langle y| \hat{K} |y\rangle \langle y \mid \psi_z\rangle$$

Now, since no photons are lost as we pass through, must have $\langle \psi_{z+\varepsilon} \mid \psi_{z+\varepsilon} \rangle = 1 = \langle \psi_z \mid \psi_z \rangle$

for all z. Then get

$$\langle \psi_{z+\varepsilon} \mid \psi_{z+\varepsilon} \rangle = \langle \psi_z \mid \psi_z \rangle + i\varepsilon \left[\langle x| \hat{K} |x\rangle - \langle x| \hat{K} |x\rangle^* \right] |\langle x \mid \psi_z\rangle|^2$$

$$+i\varepsilon \left[\langle y| \hat{K} |y\rangle - \langle y| \hat{K} |y\rangle^* \right] |\langle y \mid \psi_z\rangle|^2$$

$$+i\varepsilon \left[\langle x| \hat{K} |y\rangle - \langle x| \hat{K} |y\rangle^* \right] \langle y \mid \psi_z\rangle \langle x \mid \psi_z\rangle^*$$

$$+i\varepsilon \left[\langle y| \hat{K} |x\rangle - \langle y| \hat{K} |x\rangle^* \right] \langle x \mid \psi_z\rangle \langle y \mid \psi_z\rangle^*$$

—> must have

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

—> $\hat{K}$ is Hermitian operator.

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

From earlier discussion identify

$$\hat{U}_z = \text{ transformation operator}$$
$$\hat{K} = \text{ generator of the transformation}$$

Calculating the Transition Probability

We defined the transition probability as

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

Using

$$\hat{U}_z = e^{ik_e z}\,|e\rangle\langle e| + e^{ik_o z}\,|o\rangle\langle o| \quad , \quad |\psi_{in}\rangle = a\,|o\rangle + b\,|e\rangle$$

where $|a|^2 + |b|^2 = 1$, get

$$T(z) = \left|\langle\varphi|\left(e^{ik_e z}\,|e\rangle\langle e| + e^{ik_o z}\,|o\rangle\langle o|\right)(a\,|o\rangle + b\,|e\rangle)\right|^2$$
$$= \left|\langle\varphi|\left(be^{ik_e z}\,|e\rangle + ae^{ik_o z}\,|o\rangle\right)\right|^2 = \left|be^{ik_e z}\langle\varphi \mid e\rangle + ae^{ik_o z}\langle\varphi \mid o\rangle\right|^2$$

Now ask a specific question.

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

What is probability that it will exit as a RCP photon?

—> choose

$$|\varphi\rangle = |R\rangle = \frac{1}{\sqrt{2}}|o\rangle + \frac{i}{\sqrt{2}}|e\rangle$$

or

$$\langle\varphi \mid e\rangle = \frac{i}{\sqrt{2}} \quad , \quad \langle\varphi \mid o\rangle = \frac{1}{\sqrt{2}}$$

Then get

$$T(z) = \left|be^{ik_e z}\langle\varphi \mid e\rangle + ae^{ik_o z}\langle\varphi \mid o\rangle\right|^2 = \left|\frac{i}{\sqrt{2}}e^{ik_e z}\frac{i}{\sqrt{2}} + \frac{1}{\sqrt{2}}e^{ik_o z}\frac{1}{\sqrt{2}}\right|^2$$

$$= \frac{1}{4}\left|e^{ik_o z} - e^{ik_e z}\right|^2 = \frac{1}{4}\left(1 + 1 - e^{i(k_o - k_e)z} - e^{-i(k_o - k_e)z}\right)$$

$$= \frac{1}{2}(1 - \cos(k_o - k_e)z)$$

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

This non-trivial example clearly exhibits the power of these techniques.