Polaroids, Stern-Gerlachs and Slits - The Finale - Explanations

Remember the property of light that we called **polarization**. Let us repeat some experiments using a calcite crystal, Polaroids and a laser.

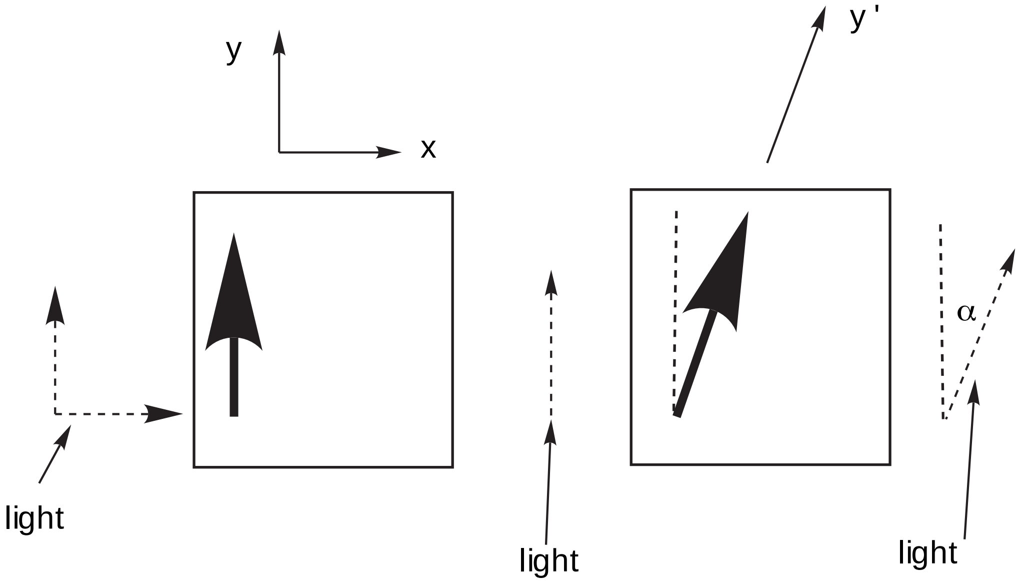
First, a calcite crystal when used in conjunction with polaroids shows definitely that are two distinct kinds of photon polarization.

Now send laser light through a Polaroid oriented so that light polarized along its x-axis passes through.

This takes the originally unpolarized beam and polarizes the light along the y-axis(d**efined by the first polaroid**). It **collapses** the polarization state as we saw earlier.

A second polaroid with its preferred axis at angle α with respect to y-axis (determined by the first polaroid) is then inserted. We choose axes x'-y' for the second polaroid. The y and y' axes make the angle α with respect to each other and the second polaroid passes light polarized in its y' direction

Brightness(intensity) measurements indicate that only a fraction $\cos^2\alpha$ of the energy of the light beam comes through the second polaroid.



This is easily seen, since the component of the vector parallel to the y' axis is a component = $\cos\alpha$ times the original vector length(=1 for convenience) in x direction. Since the intensity(brightness) of the light is proportional to the square of the $\vec{E}$ vector(length squared), we end up with $\cos^2\alpha$ of the light (intensity) energy that passes through the first polaroid subsequently passing through the second polaroid.

Now in the language of state vectors and operators we have:

$$|y\rangle = |\text{photon state after 1st polaroid}\rangle = \hat{P}_y|\psi\rangle$$

where $\hat{P}_y$ is the y-polaroid operator.

If

$$|\psi\rangle = \frac{1}{\sqrt{2}}|x\rangle + \frac{1}{\sqrt{2}}|y\rangle \quad \text{when written in the } |\mathrm{x}\rangle,|y\rangle \text{ basis}$$

then

$$\hat{P}_y = |y\rangle\langle y|$$

The $\frac{1}{\sqrt{2}}$ factor simply says that the number of photons getting through is 1/2 of the total .

At this point we choose the $|\mathrm{x}'\rangle,|y'\rangle$ basis and write

$$|x\rangle = \sin\alpha|x'\rangle + \cos\alpha|y'\rangle$$

Therefore, after the first polaroid we are in the state $|x\rangle = \sin\alpha|x'\rangle + \cos\alpha|y'\rangle$. Now

$$y'-polaroid = \hat{P}_{y'} = |y'\rangle\langle y'|$$

so that the number of photons is reduced by a further factor of $\cos^2\alpha$.

Or in words, we start in a laser state, which after the first polaroid is a state polarized in the y direction. This is equivalent to a state which is a linear combination of a x'-polarized state and a y'-polarized states with respect to the axes of the second polaroid with weights $\sin\alpha$ and $\cos\alpha$ respectively.

Due to the orientation of the second polaroid, only the y'- polarized state of the light after the first polaroid can get through and so on.

This explanation fits well, if we are thinking of light waves and their intensities(large numbers of photons).

So, we now repeat the experiment, but we first reduce the intensity of the light such that only one photon is in the vicinity of the material at any time.

This photon really has a problem!

In order for the new experiment to agree with the old one, this single photon must allow only a fraction $\cos^2\alpha$ of its energy get through the second polaroid!

The problem is that the photon is not divisible! The photon can do only one of two things:

either it all gets through (no absorption of the photon)
or
none of it gets through (total absorption of the photon).

Thus, we either get **1 or 0** photon as the result of the experiment.

If we want to get agreement with the wave results then we must assume that the photon sometimes gets through and sometimes does not!

We cannot predict what any particular photon will do, but if we watch a large enough number than we will find that a fraction $\cos^2\alpha$ have gotten through. We thus are led to assign a probability $\cos^2\alpha$ that the photon is transmitted and a probability $\sin^2\alpha$ that the photon is absorbed.

Note that the sum of the probabilities for all possible outcomes = 1 ($\cos^2\alpha+\sin^2\alpha=1$) since something must happen. Suddenly, we have radical unpredictability of individual events appearing in quantum mechanics.

What else has happened in this experiment?

The direction of polarization or the polarization state, has changed during the experiment!

After first polaroid the light was in a linear combination of x' and y' polarization states (with respect to the second Polaroid) and when it came out of second Polaroid it was in an y'-polarization state!

What does this mean for single photons?

Only one interpretation seems possible -- that the polarization state of the photon was **instantaneously** changed(collapsed) in the second polaroid from a mixed (superposition) polarization state to a y'-polarized state if it got through.

Let us summarize these results in the proper language of quantum mechanics.

Any state can be considered a linear combination or superposition of other states (i.e., x and y polarization states) where the other states that make up an arbitrary state are all mutually exclusive, i.e., there is no amount of x-polarized state in a y-polarized state and vice versa(this is what we mean by an orthonormal basis).

In the correct notation it looks like:

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

where $|a|^2 + |b|^2 = 1$ and $|a|^2$ is the probability that the photon in state $|p\rangle$ will be transmitted in the x-oriented polaroid and $|b|^2$ is the probability that it will be absorbed.

What has this idea done to the world?

Do any states have definite properties?

What about $|x\rangle$ or $|y\rangle$?

If we observe that there are several outcomes from some experiment on a beam of light, then how do we interpret this result in terms of the initial state of the beam?

Do we say the original state was a linear combination of states representing all the possibilities as with x and y polarization?

That is exactly what we are forced to say because light can act like a photon (like a particle).

The superposition principle has taken us straight to the heart of the elusive indeterminacy of quantum mechanics. We now have a way to describe the "**state**" of an electron when it is "**prepared**" in an experiment.

These states will be the vectors we have been describing.

Any experiment, however, involves a second part, namely, the act of measurement -- like determining where the electron is or how it is moving.

Measurement is represented by operators and some operators are incompatible.

We defined incompatibility by saying that we could not simultaneously know both observables. It also means that the order of measurement matters (remember the commutator).

Consider the following experiment using the polaroids. We use two polaroids oriented at an angle α with respect to each other, call them $\hat{P}_1$ and $\hat{P}_2$.

(1) Initial state of a photon = $|\text{after calcite}\rangle$ → 2 separated beams of orthogonal polarizations

(2) Measure polarizations with $\hat{P}_1$ and $\hat{P}_2$ in the order shown below

$$\hat{P}_1\hat{P}_2|\text{after calcite}\rangle$$

We adjust the first polaroid(2) so that one of the beams disappears.

The final result is only one beam because the other is extinguished by the first measurement!

But if we reverse the order of measurement(without changing the polaroid orientations) we have the operation

$$\hat{P}_2\hat{P}_1|\text{after calcite}\rangle$$

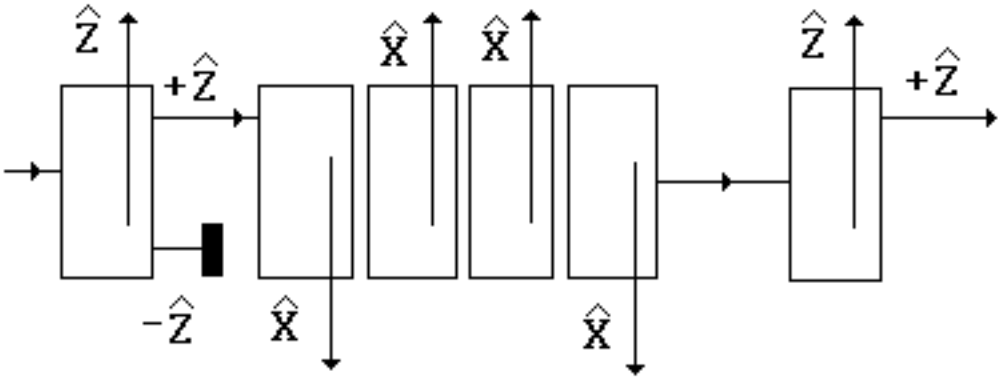
In this case both beams are once again present at the end.

The important result is that none gets through in one order and some gets through in the reversed order which says that $\hat{P}_1$ and $\hat{P}_2$ are incompatible and that they are **not simultaneously measurable.**

Thus for incompatible observables the order of measurements must be carefully monitored. That is the real content of the Heisenberg uncertainty principle.

Now let us return to the Stern-Gerlach Apparatus

Remember Experiment #5 from earlier. It is shown below.



In this experiment a beam enters the $SG\hat{\mathbf{z}}$ device and we block the exiting $\hat{\mathbf{J}}\bullet\hat{\mathbf{z}}=-\frac{\hbar}{2}$ beam. The $\hat{\mathbf{J}}\bullet\hat{\mathbf{z}}=+\frac{\hbar}{2}$ beam is sent into the modified $SG\hat{\mathbf{x}}$ device and the exit beam is sent into the final $SG\hat{\mathbf{z}}$ device. We **DO NOT** block any of the paths in the modified $SG\hat{\mathbf{x}}$ device. Since the beam entering the modified $SG\hat{\mathbf{x}}$ device is reconstructed before it exits (we already saw that it does this in the Experiment #4 earlier), we are **NOT** making a measurement of $\hat{\mathbf{J}}\bullet\hat{\mathbf{x}}$ using the modified $SG\hat{\mathbf{x}}$ device as we did when we used the original $SG\hat{\mathbf{x}}$ device in Experiment #3 earlier.

Now we send in N particles and N/2 are in the $\hat{\mathbf{J}}\bullet\hat{\mathbf{z}}=+\frac{\hbar}{2}$ beam as before.

However, now, instead of finding N/8 particles in the final $\hat{\mathbf{J}}\bullet\hat{\mathbf{z}}=+\frac{\hbar}{2}$ beam, we find N/2 particles. **ALL** the particles make it through unchanged, even though there are $SG\hat{\mathbf{x}}$ devices in between. It behaves as if the modified $SG\hat{\mathbf{x}}$ device was not there at all and nothing is randomizing the z information.

We might have assumed that 50% of the particles in the $\hat{\mathbf{J}}\bullet\hat{\mathbf{z}}=+\frac{\hbar}{2}$ beam before the modified $SG\hat{\mathbf{x}}$ device would emerge in the $\hat{\mathbf{J}}\bullet\hat{x}=+\frac{\hbar}{2}$ state and the other 50% would emerge in the $\hat{\mathbf{J}}\bullet\hat{x}=-\frac{\hbar}{2}$ state.

Experiment #5 says this cannot be true, since if it were true, then we would have two beams coming out of the final $SG\hat{\mathbf{z}}$ device, each with 50% of the particles.

Our results are **incompatible** with the statement that the particles passing through the modified $SG\hat{x}$ device are **either** in the state $\hat{\mathbf{J}}\bullet\hat{x}=+\frac{\hbar}{2}$ **or** the in the state $\hat{\mathbf{J}}\bullet\hat{x}=-\frac{\hbar}{2}$.

In fact, if we carry out this experiment with a very low intensity beam where only one particle at a time is passing through the apparatus, then we observe that **each** particle emerging from the final $SG\hat{\mathbf{z}}$ device is in the state $\hat{\mathbf{J}}\bullet\hat{\mathbf{z}}=+\frac{\hbar}{2}$. This eliminates any explanation of the result that would invoke interactions among the particles while they are in the apparatus.

For beams of particles, we have been talking in terms of percentages or fractions of the particles as experimental results.

For a single particle, however, it is not possible to predict with certainty the outcome of a measurement in advance. We have seen this earlier in our experiments with photons polaroids. We are **only** able to use probability arguments in this case.

Probabilities alone, however, are not enough to explain Experiment #5. We came to the incorrect conclusion because we made a mistake.

There are two ways the photon could get through the apparatus and these ways are **indistinguishable**.

When we calculate a probability, i.e., $|\langle a|b\rangle|^2$ = the probability that we will find the system in the state $|a\rangle$ after a measurement if before the measurement it was in the state $|b\rangle$. We define the quantity

$\langle a|b\rangle$ = the **probability amplitude** that we will find the system in the state $|a\rangle$ after a measurement if before the measurement it was in the state $|b\rangle$

We added the separate probabilities of the indistinguishable ways to get the total probability, whereas the correct result is to add the amplitudes of the indistinguishable ways and then square the total amplitude to get the correct result.

We eliminated interference effects and got the incorrect result!!

When we don't actually make a measurement as in the modified $SG\hat{\mathbf{x}}$ device, we must add amplitudes and not probabilities.

We can now use our formalism, which should allow us to explain the experiments correctly.

We need a 2-dimensional vector space to describe these physical systems.

As a basis for this space we can use any of the sets $\{|+\hat{\mathbf{n}}\rangle, |-\hat{\mathbf{n}}\rangle\}$ corresponding to the definite values $\pm\frac{\hbar}{2}$ for $\hat{\mathbf{J}} \bullet \hat{\mathbf{n}}$. Each of these is an orthonormal basis where

$$\langle +\hat{\mathbf{n}}|+\hat{\mathbf{n}}\rangle = 1 = \langle -\hat{\mathbf{n}}|-\hat{\mathbf{n}}\rangle \quad \text{and} \quad \langle -\hat{\mathbf{n}}|+\hat{\mathbf{n}}\rangle = 0$$

Any arbitrary state can be written as a superposition of the basis states

$$|\psi\rangle = \langle +\hat{\mathbf{n}}|\psi\rangle |+\hat{\mathbf{n}}\rangle + \langle -\hat{\mathbf{n}}|\psi\rangle |-\hat{\mathbf{n}}\rangle$$

and the operators can be written as

$$\hat{\mathbf{J}} \bullet \hat{\mathbf{n}} = \frac{\hbar}{2}|+\hat{\mathbf{n}}\rangle\langle +\hat{\mathbf{n}}| - \frac{\hbar}{2}|-\hat{\mathbf{n}}\rangle\langle -\hat{\mathbf{n}}| = J_n$$

Expectation values are

$$\langle \hat{J}_z \rangle = \langle \psi|\hat{J}_z|\psi\rangle = \frac{\hbar}{2}|\langle +\hat{\mathbf{z}}|\psi\rangle|^2 - \frac{\hbar}{2}|\langle -\hat{\mathbf{z}}|\psi\rangle|^2$$

The quantity

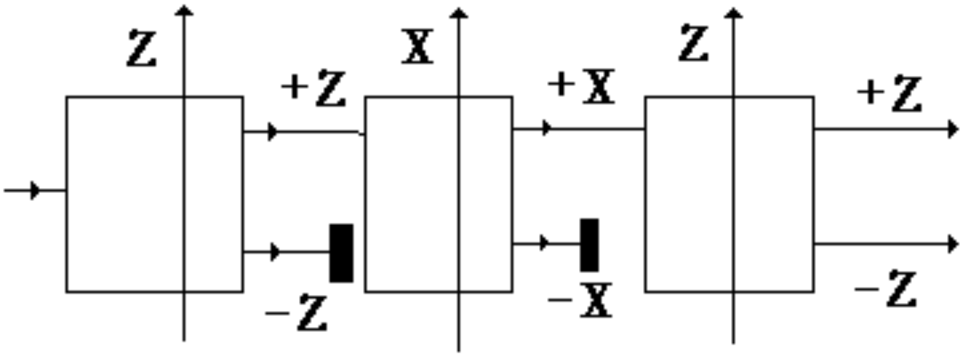
$$\langle +\hat{\mathbf{n}}|\psi\rangle = \text{probability amplitude for measurement of } |+\hat{\mathbf{n}}\rangle \text{ when in state } |\psi\rangle$$

and

$$\left|\langle +\hat{\mathbf{n}}|\psi\rangle\right|^2 = \text{corresponding probability}$$

Let us Analyze Experiment #3 in Detail

The set up is shown below:



In Experiment #3, the state before entering the first $SG\hat{\mathbf{z}}$ device is

$$|\psi_1\rangle = a|+\hat{\mathbf{z}}\rangle + b|-\hat{\mathbf{z}}\rangle \quad \text{where} \quad |a|^2 + |b|^2 = 1$$

Since the $SG\hat{\mathbf{z}}$ device is a measurement of $\hat{J}_z$, after the $SG\hat{\mathbf{z}}$ device the state is $|\psi_2\rangle = |+\hat{\mathbf{z}}\rangle$ (remember the we blocked the $|-\hat{\mathbf{z}}\rangle$ path).

It is very important to realize that we cannot answer a question about a measurement unless we express the state in the basis consisting of the eigenvectors of the operator representing the observable being measured.

Remember we call this the home space of the observable.

That is why we used the $\hat{J}_z$ eigenvectors to discuss the measurement made using the $SG\hat{\mathbf{z}}$ device.

Now, we are going to make a measurement of $\hat{J}_x$ in the $SG\hat{\mathbf{x}}$ device. So we now switch to a basis consisting of the $\hat{J}_z$ eigenvectors. We can write

$$|+\hat{\mathbf{x}}\rangle = \langle +\hat{\mathbf{z}}|+\hat{\mathbf{x}}\rangle|+\hat{\mathbf{z}}\rangle + \langle -\hat{\mathbf{z}}|+\hat{\mathbf{x}}\rangle|-\hat{\mathbf{z}}\rangle$$

One of our experiments tells us that when we send a particle in a $|+\hat{\mathbf{x}}\rangle$ state through an $SG\hat{\mathbf{z}}$ device, the probability = 1/2 that we find $|+\hat{\mathbf{z}}\rangle$ and 1/2 that we find $|-\hat{\mathbf{z}}\rangle$. This means that

$$\left|\langle +\hat{\mathbf{z}}|+\hat{\mathbf{x}}\rangle\right|^2 = \left|\langle -\hat{\mathbf{z}}|+\hat{\mathbf{x}}\rangle\right|^2 = \frac{1}{2}$$

or

$$\langle +\hat{\mathbf{z}}|+\hat{\mathbf{x}}\rangle = \frac{1}{\sqrt{2}} \quad \text{and} \quad \langle -\hat{\mathbf{z}}|+\hat{\mathbf{x}}\rangle = \frac{1}{\sqrt{2}}$$

and

$$|+\hat{\mathbf{x}}\rangle = \frac{1}{\sqrt{2}}|+\hat{\mathbf{z}}\rangle + \frac{1}{\sqrt{2}}|-\hat{\mathbf{z}}\rangle$$

Similarly,

$$|-\hat{\mathbf{x}}\rangle = \frac{1}{\sqrt{2}}|+\hat{\mathbf{z}}\rangle - \frac{1}{\sqrt{2}}|-\hat{\mathbf{z}}\rangle$$

which explains why we see the two beams of equal intensity at the end. Thus, our formalism easily handles Experiment #3.

Now to finish Experiment #5

Since we are not actually measuring $\hat{J}_z$ in the center apparatus, the incoming $|+\hat{\mathbf{z}}\rangle$, which is a superposition of $|+\hat{\mathbf{x}}\rangle$ and $|-\hat{\mathbf{x}}\rangle$ has an amplitude for passing through as $|+\hat{\mathbf{x}}\rangle$ and amplitude for passing through as $|-\hat{\mathbf{x}}\rangle$.

The probability amplitude of observing a $|-\hat{\mathbf{z}}\rangle$ at the end is

$$\langle -\hat{z}|+\hat{z}\rangle = 0 = \langle -\hat{z}|+\hat{x}\rangle\langle +\hat{x}|+\hat{z}\rangle + \langle -\hat{z}|-\hat{x}\rangle\langle -\hat{x}|+\hat{z}\rangle = 0$$

where

$$\langle -\hat{z}|+\hat{x}\rangle\langle +\hat{x}|+\hat{z}\rangle = \text{amplitude for passing through as } |+\hat{x}\rangle = \frac{1}{2}$$

$$\langle -\hat{z}|-\hat{x}\rangle\langle -\hat{x}|+\hat{z}\rangle = \text{amplitude for passing through as } |-\hat{x}\rangle = -\frac{1}{2}$$

So the probability is correctly = 0.

Our mistaken calculation goes as follows. Instead of adding the amplitudes and the squaring, we squared the amplitudes and then added, as if these were distinguishable processes. We obtained

$$\left|\frac{1}{2}\right|^2 + \left|-\frac{1}{2}\right|^2 = \frac{1}{2} \rightarrow 50\%\ are\ |-\hat{z}\rangle$$

Therefore, we have the most powerful rule of quantum theory.

If a process can happen in several indistinguishable ways, the correct probability for the process is obtained by adding the amplitudes for all the indistinguishable ways and squaring the total amplitude.

In this way we retain interference effect and see quantum behavior.

Squaring the individual amplitudes and adding removes all interference effect and gives erroneous classical type results.

In the color-hardness experiment. The correct answer is the think of the magenta electron as passing through the apparatus as a superposition of hard and soft. Add the amplitudes for each way and then square to get the correct probabilities. If we look to see what path they are on, the paths are no longer indistinguishable and we add probabilities (not amplitudes) and get a different result.

Let us see this work again.

Two slit experiment:

The only answer will be that the electron state is a superposition of states corresponding to passing through both slits . We have

$$
\begin{aligned}
&\left|\text{initial state}\right\rangle \\
&\left|\text{entering slits region}\right\rangle = \hat{\mathrm{T}}\left|\text{initial state}\right\rangle \\
&\left|\text{leaving slits region}\right\rangle = \hat{\mathrm{S}}\left|\text{initial state}\right\rangle \quad , \quad \hat{\mathrm{S}} = \text{slit operator} \\
&\qquad = a\left|\text{passed through slit 1}\right\rangle + b\left|\text{passed through slit 2 }\right\rangle \\
&\left|\text{at screen detectors}\right\rangle = \hat{\mathrm{T}}\left|\text{leaving slits region}\right\rangle \\
&\qquad = \hat{\mathrm{T}}\left(a\left|\text{passed through slit 1}\right\rangle + b\left|\text{passed through slit 2 }\right\rangle\right) \\
&\qquad = \sum_k a_k\left|\mathrm{k}^{\mathrm{th}}\text{ detector}\right\rangle
\end{aligned}
$$

where

$$\left|a_k\right|^2 = \text{probability of detection at }\mathrm{k}^{\mathrm{th}}\text{ detector}$$

and

$\left|\mathrm{k}^{\mathrm{th}}\text{ detector}\right\rangle$ = state of being at k^{th} detector

We are measuring a position variable in this case.

Now, a_k will have contributions from both slits,.i.e.,

$$a_k = a_k(1) + a_k(2)$$

Then, since the paths are indistinguishable,

$$|a_k|^2 = |a_k(1) + a_k(2)|^2 \neq |a_k(1)|^2 + |a_k(2)|^2$$

which allows for interference (cancellation) only if both slits are present.

If you try to detect which slit the electron went through, then you collapse the state into one or the other and the interference pattern must go away and it does!!!

That is, if we try to detect if the electron passes through slit #2 then we have

if it does ---> $a_k = a_k(2)$

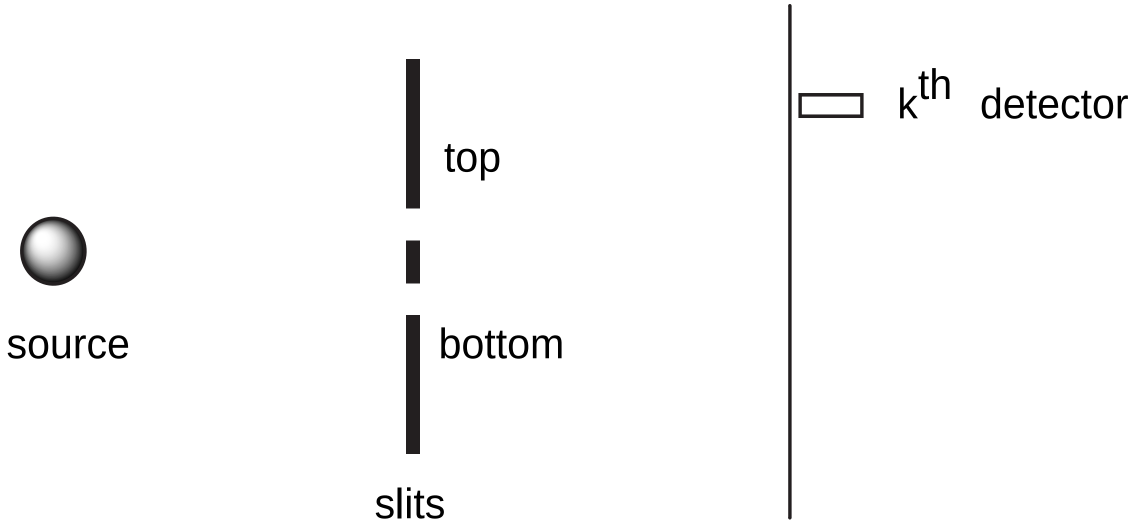
if it does not --> it passes through slit #1 ---> $a_k = a_k(1)$

In both cases the interference disappears.

From the detector probabilities we cannot go backwards in time and reconstruct the paths and find out which slits the electrons went through! In classical physics we could!!

Saying the same thing slightly differently (because it is so important)

Two slits ......



$$|\psi_{after-slits}\rangle = \hat{T}|\psi_{initial}\rangle = a_{top}|\psi_{top}\rangle + a_{bottom}|\psi_{bottom}\rangle$$

i.e., a **superposition of all the possible ways** for the result to occur.....

Our probability discussions then say that

$$\begin{aligned} A_k &= \text{probability amplitude to be detected at k}^{\text{th}}\text{ detector} \\ &= \langle k | \psi_{after-slits} \rangle = \langle k | \left(a_{top} | \psi_{top} \rangle + a_{bottom} | \psi_{bottom} \rangle \right) \\ &= a_{top} \langle k | \psi_{top} \rangle + a_{bottom} \langle k | \psi_{bottom} \rangle \\ &= A_{top} + A_{bottom} \\ &= \text{sum of amplitudes for all indistinguishable ways} \end{aligned}$$

Finally,

$$\begin{aligned} P_k &= \text{probability to be detected at k}^{\text{th}}\text{ detector} \\ &= \left| \langle k | \psi_{after-slits} \rangle \right|^2 \\ &= \left| A_{top} + A_{bottom} \right|^2 \end{aligned}$$

which gives the correct interference pattern

Another polarization example

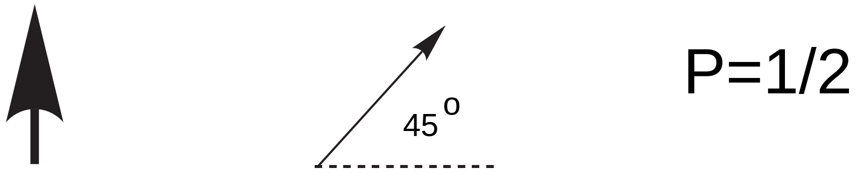
$\Uparrow \quad \rightarrow \quad \Pr ob = 0 \qquad \Uparrow \quad \uparrow \quad \Pr ob = 1$

$\Rightarrow \quad \rightarrow \quad \Pr ob = 1 \qquad \Rightarrow \quad \uparrow \quad \Pr ob = 0$

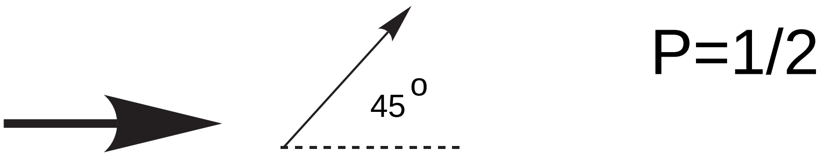
GO-NOGO processes

where $\Uparrow \quad \rightarrow$ are preferred and polarization directions, respectively.

Two possible polarization states.....

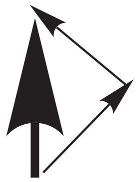


$\Uparrow = |90°\rangle$ passing through

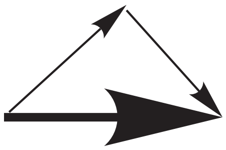


$\Rightarrow = |0°\rangle$ passing through

Using a different basis we can write



$$\Uparrow = |90^\circ\rangle = \frac{1}{\sqrt{2}}|45^\circ\rangle + \frac{1}{\sqrt{2}}|135^\circ\rangle$$



$$\Rightarrow = |0^\circ\rangle = \frac{1}{\sqrt{2}}|45^\circ\rangle - \frac{1}{\sqrt{2}}|135^\circ\rangle$$

which explains why the probabilities are 1/2.

Now suppose a beam composed of 90° and 0° photons comes to a 45° polaroid. For each beam, 1/2 of photons get through. After getting through they are all polarized at 45°.

The important point is we cannot tell which beam a particular 45° photon originally came from .... we have **erased that information**!!!

More about this point later.