Start of Part #2

No theory yet - This story I will now tell will help push us in the correct direction!

A Simple Introduction to Quantum Ideas via Experiments

I am going to tell you a story about a world of electrons that have 2 properties

Color and Hardness

In this imagined Microworld:

Electrons have 2 properties ==> color and hardness

Note that In the Real world:

Electrons also have 2 properties ==> z- and x-spin (orthogonal directions)

and

Photons also have 2 properties => horizontal and vertical polarization

The Strange World of Color and Hardness

Known experimental facts:

Color has **two** possible values = green or magenta.

This means that **nothing else** has ever been **measured** for electron color.

Similarly, for hardness property.

Electrons are either hard or soft (**only** possibilities).

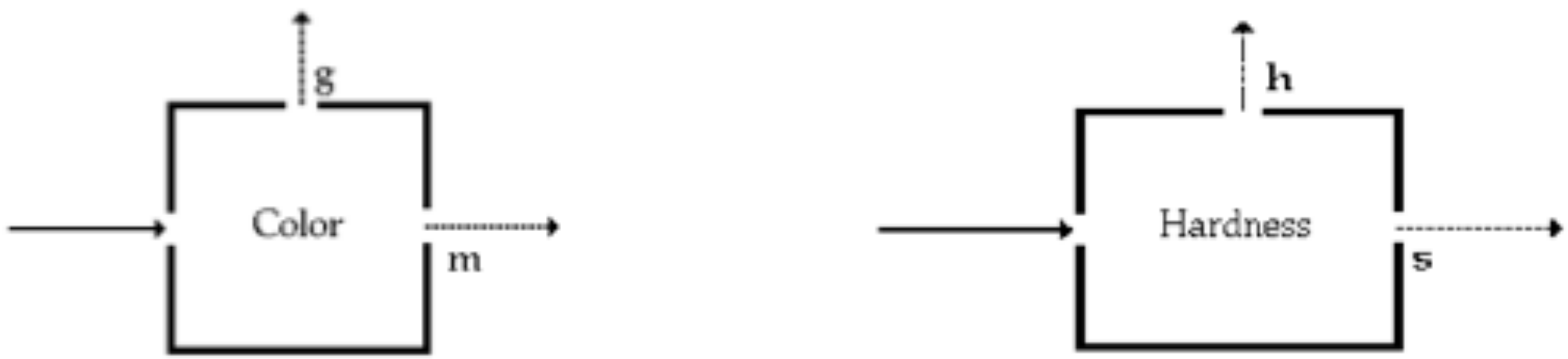
Nothing else ever **measured**.

Color/Hardness Boxes (have 3 apertures)

Electrons enter boxes through **left** aperture and exit through one of other two apertures.

These boxes represent the same physics as real world devices but without complications.

[Examples : Stern-Gerlach devices and Polarizers as we will see later.]

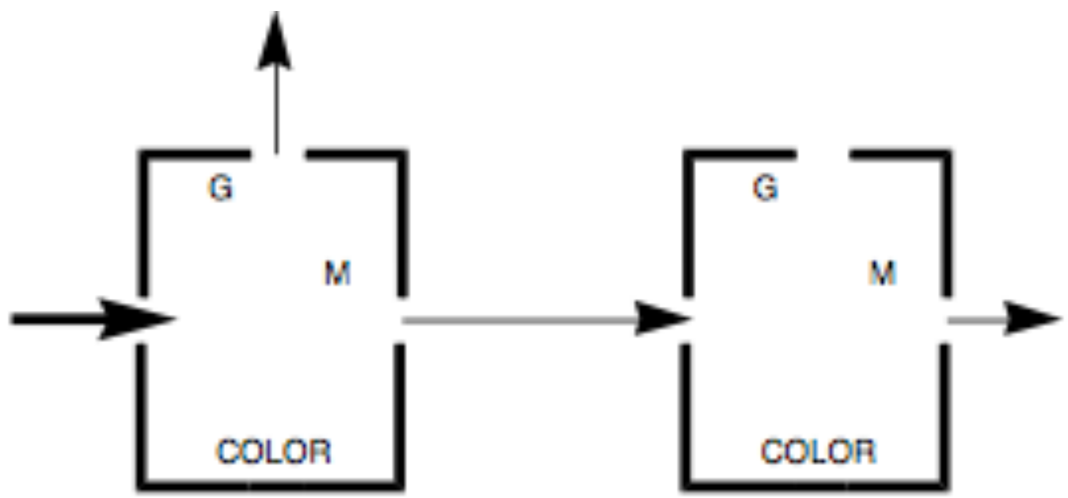


Boxes **separate** electrons in physical space (put them onto **different paths**) according to the value of color or hardness as shown.

Thus, in either case, one can distinguish the electrons **after** they pass through a box by their final position(path) in real physical space, i.e., we have **separate** beams at end.

AN EXPERIMENTAL PROPERTY

Measurements with hardness or color boxes are **repeatable**.



The 1st color box determines the electron color and we choose the magenta beam.

Then, without tampering(must define this concept) between measurements the magenta beam is subsequently (immediately=> $\Delta t=0$) sent into left aperture of another color box.

The electron will (with probability = 1) emerge from 2nd box through magenta aperture.

This **same** property of repeatability holds for green electrons and color boxes

and for hard or soft electrons with hardness boxes.

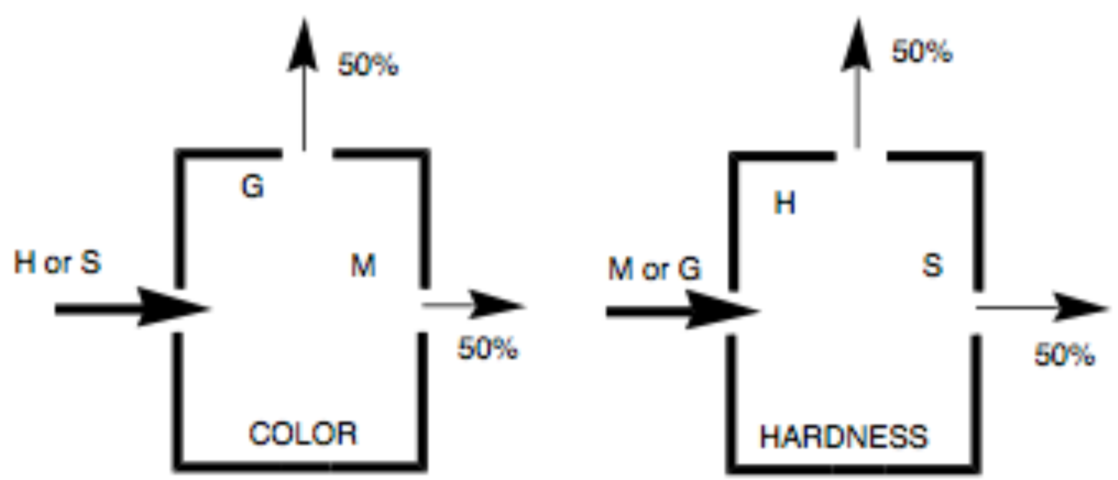
It is **simply** a statement of the way the real microworld and its measuring devices work in the laboratory.

So our representation of the real world should behave that way also.

Suppose we suspect the possibility that color and hardness properties are related in some way.

We must investigate => check for **correlations** (or **relationships**) between the measured values of hardness and color properties.

The boxes allow for checking whether such correlations exist.



No such correlations exist,

i.e., of any large collection of, say, magenta electrons,

all fed into left aperture of hardness box, **precisely** 1/2 emerge through hard aperture,

and **precisely** 1/2 emerge through soft aperture - one cannot predict which goes where!

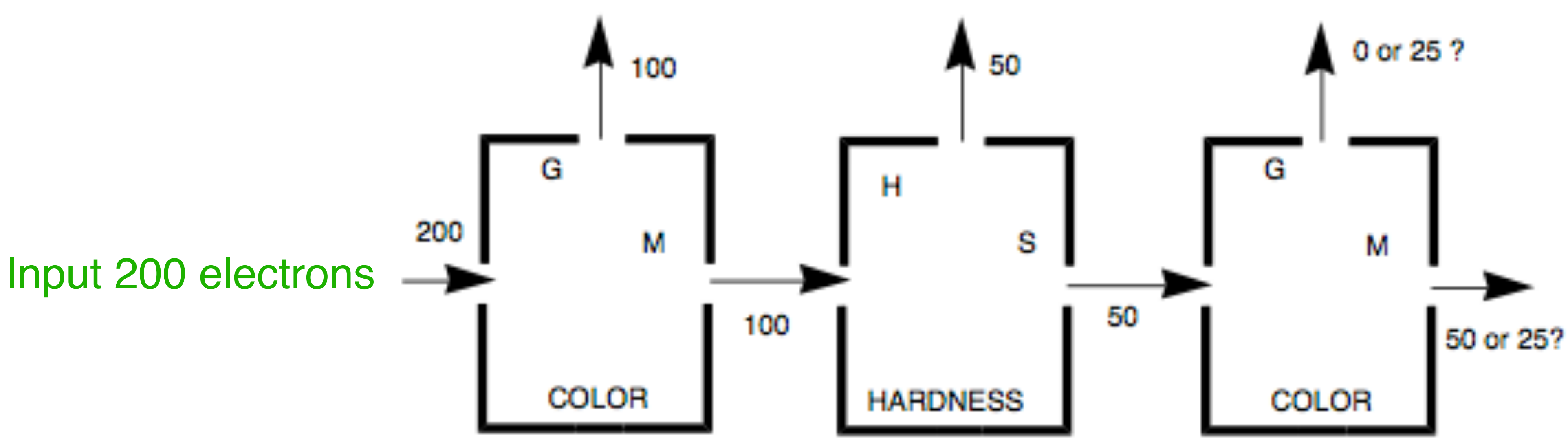
Same result occurs for green electrons and similar results hold if we send hard or soft electrons into color box - all results are exactly 50% —> **there are no correlations**.

Color (hardness) of electron apparently gives no information about hardness (color).

This usually means that no correlations seem to exist.

Now, suppose we set up a sequence of **three** boxes.

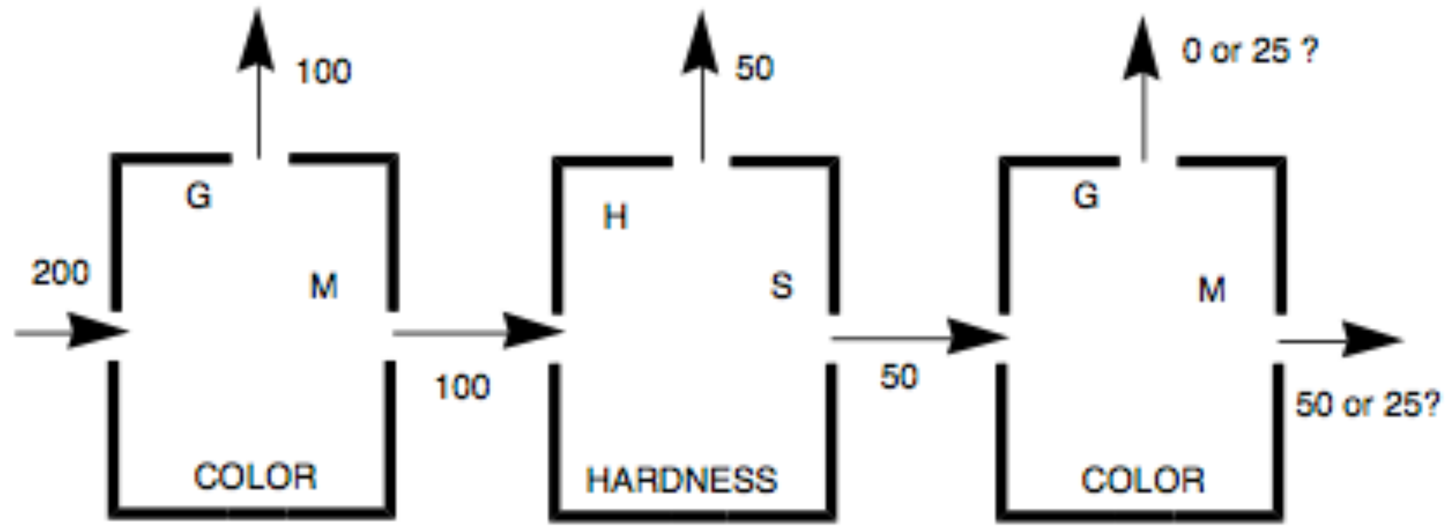
Color box, then hardness box and finally color box.



Suppose an electron(1 of 100) that emerges from magenta aperture of 1st color box

then goes into left aperture of hardness box

(no tampering => no other measurements or no time interval).



Then suppose an electron that emerges(50 out of 100 do) from hardness box through soft aperture (1/2 of input 100 magenta electrons)

and are sent into left aperture of last color box

(again no measurements or time intervals allowed).

Presumably, the electron(1 out of 50) that enters that last box

is known to be **BOTH** magenta **and** soft,

which were results of two previous measurements just made on it!

If so, then we expect the electrons(50) to emerge from magenta aperture(probability = 1),

confirming the result of 1st measurement.

Any **reputable** classical physicist would say at this point

that electrons entering last color box are known to be **magenta/soft**.

They have properties and we are just measuring again what they already have!!!!

Repeating, in the classical world, particles have **objective reality** - have **real properties**

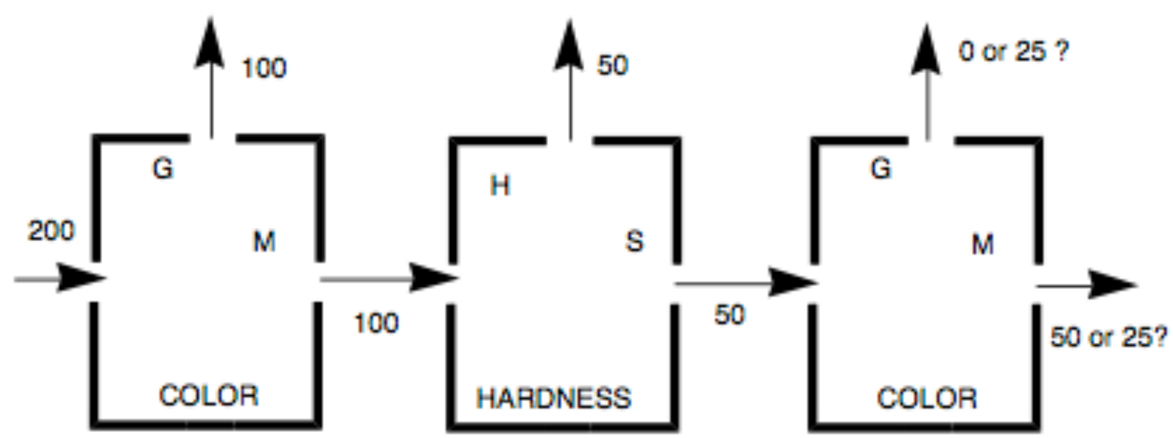
and we just **measure to find out** values of properties.

We just used 2 experiments to determine these properties!

PROBLEM: **this is not what happens in the real microworld in this experiment!**

Precisely half of electrons entering last box emerge from magenta aperture

and precisely half emerge from green aperture!!



Therein lies the fundamental puzzle of the quantum world.

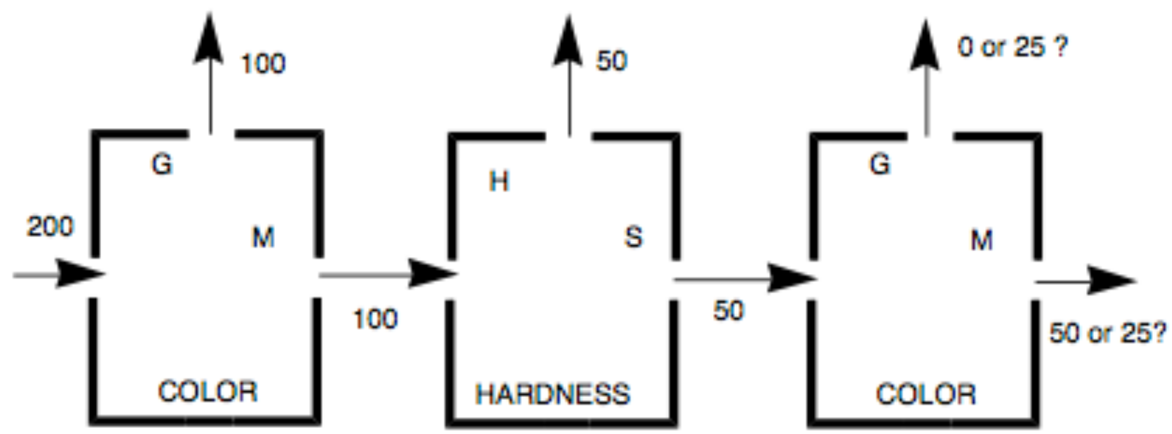
Any theory we develop must account for this result.

In fact, if 1st two measurements give magenta/soft, magenta/hard, green/soft or green/hard

which represents all possible cases, then,

when any of these beams sent into last box,

precisely half emerge from each aperture!!



It seems as if presence of hardness box between two color boxes

constitutes some sort of color measurement or color tampering and vice versa.

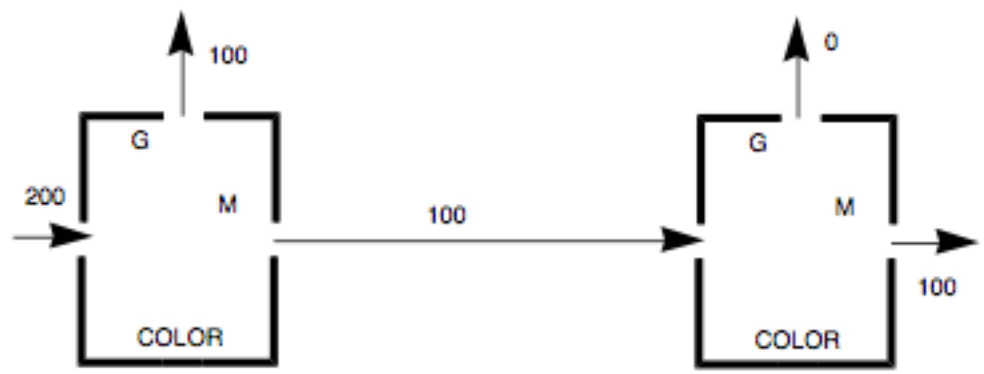
Hardness box **seems to be changing** half of magenta electrons into green electrons.

As we will see, this will turn out to be an **incorrect way** of thinking about the results of the experiment!

Hardness box must be the blame since if it were not there,

then the last box would only see magenta electrons

(a **different** experiment that corresponds to repeatability)



This all seems so weird (non-classical)

—> we must question (challenge) all features of experiment before we accept story as truth.

Perhaps hardness box is poorly built (not enough $$$$ from Congress in our research grant).

Maybe measuring hardness correctly disrupts color —> bad design.

Two fundamental questions arise:

1. Can hardness boxes be built that will measure hardness without disrupting color?

2. In case of poorly built apparatus, it seems as if half of electrons change color….

   What is it that determines which electrons have color changed and which do not change?

Question 2 can be answered by checking correlations —> **none exist**.

Those electrons that have color changed by hardness box

and those electrons whose color is not changed by hardness box

do not differ from one another in any **measurable** way.

So 2nd question has no answer that we can figure out from measurements.

If we believe, as I do, that **I can only know properties that I can measure**,

then it means there is NO answer to this question,

i.e., there is no property of electrons that determines which electrons get color changed.

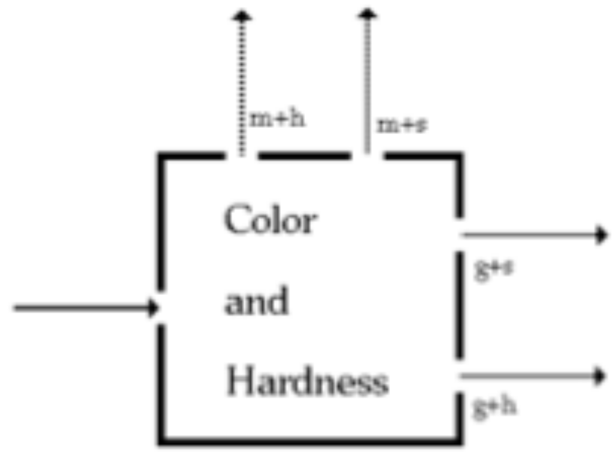
This is completely counter to classical notion of cause/effect and objective reality!!!

Question 1: no matter how we build hardness box (remember a device qualifies as a hardness box if it can separate electrons by hardness value) **all disrupt color measurements** - all change half of entering electrons.

Any hardness (color) measurement seems to **randomize** next color (hardness) measurement, i.e., make it 50% green/ 50% magenta.

What about building a color AND hardness box?

This would need 5 apertures as shown.



Has to consist of hardness boxes and color boxes or their equivalents

because we must be measuring hardness and color in some way.

But as we have already seen, whichever box the electrons pass through **last**

provides reliable information ONLY about that measured quantity

and the other quantity is **randomized** (i.e., half/half).

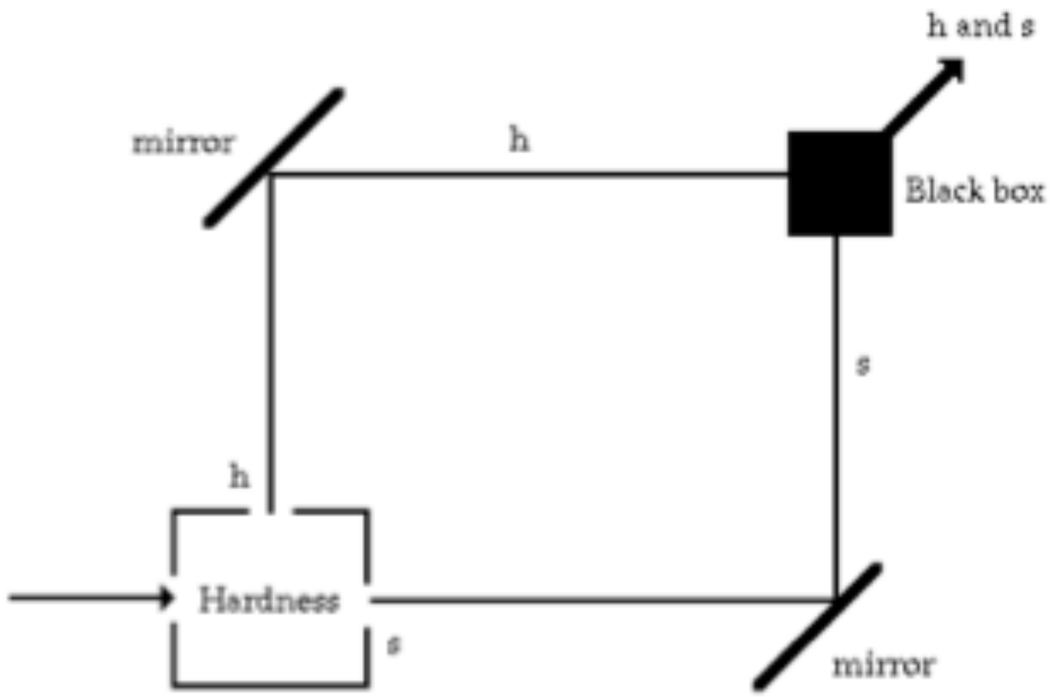
No one has succeeded in building device

which can simultaneously measure both color and hardness.

Seems fundamentally beyond our means no matter how clever we are.

We will find the reason later!

Probing Deeper



We now build a more complicated experimental setup involving two paths as shown.

We will see this device often later - it is called a Mach-Zehnder Interferometer.

The mirrors are just reflectors which do not affect color/hardness values.

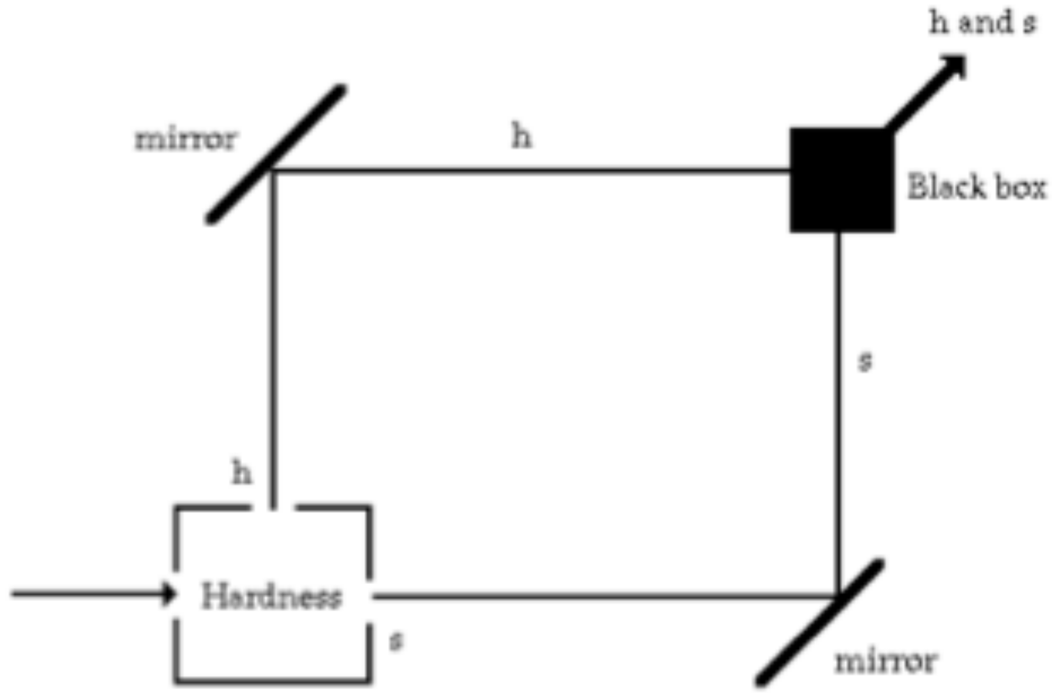
That means it has been experimentally checked to be true.

In this device, the hard and soft electrons follow **different paths in physical space**

and eventually the electrons are recombined into single beam again in black box at end.

All the black box does is recombine two beams into one beam by direction changes (like light beams with mirrors)

Again, this is done without changing hardness or color values (checked experimentally).



So if we start

with a mixed (hard + soft) beam of electrons,

entering the first hardness box

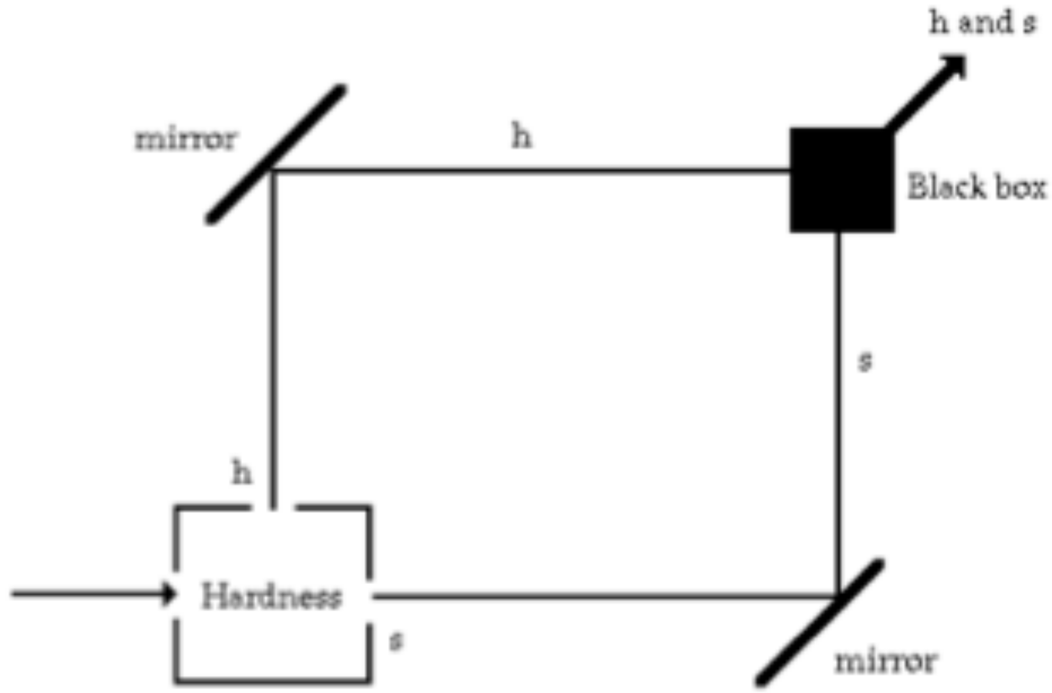
then we end up

with a mixed (hard + soft) beam (unchanged) at the end.

The effectiveness of device has been checked separately for both hard and soft electrons and it works.

Start with hard —> hard at end and they **went** on hard path!

Start with soft —> soft at end and they **went** on the soft path!



This is very important so I repeat:

If hard or soft electrons are sent in separately,

they simply travel along different paths and end up in same place

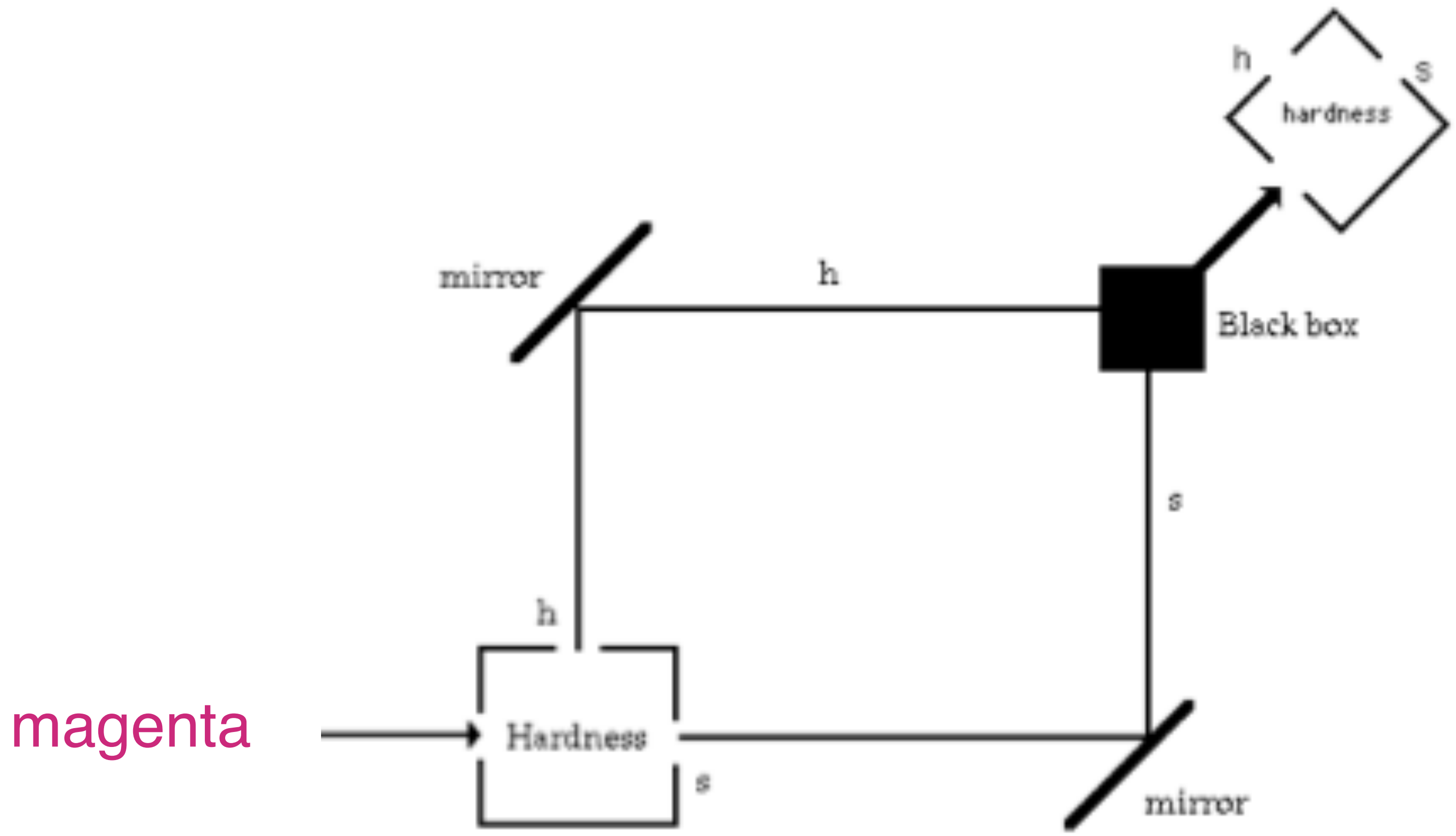
with their hardness property **unchanged**.

They go on their separate paths and emerged unchanged at end

Now some experiments can be done with apparatus.

All of these experiments have **actually** been done in real laboratories with equivalent setups.

Listen carefully now to see where your classical mind is misleading you.



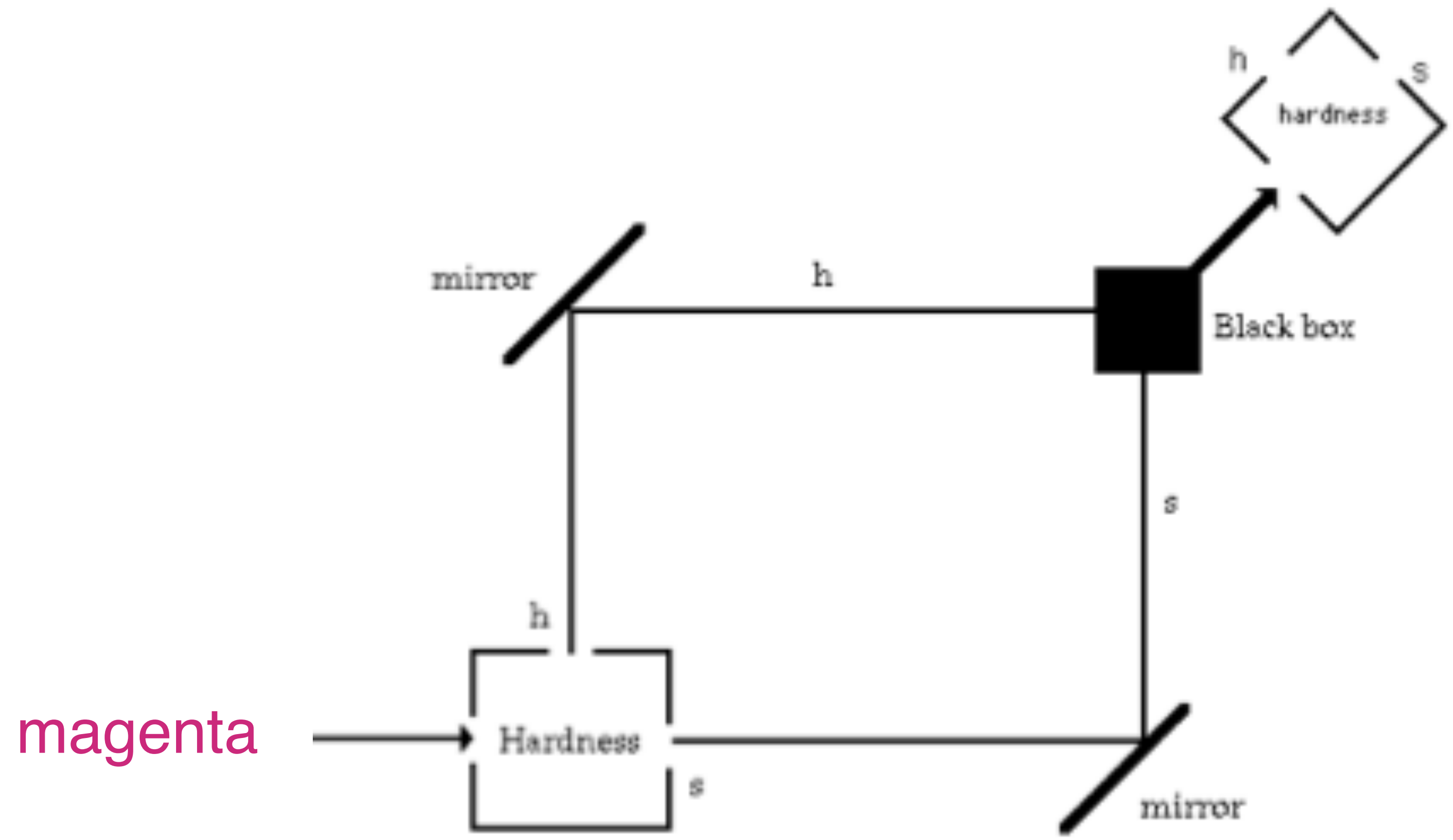
Send magenta electrons into 1st hardness box.

At end (after recombination of beams)

add a final hardness box

and thus we are measuring hardness at that point

i.e., after the “magenta” electrons have **“passed”** through the apparatus.



Analysis:

For entering magenta electrons,

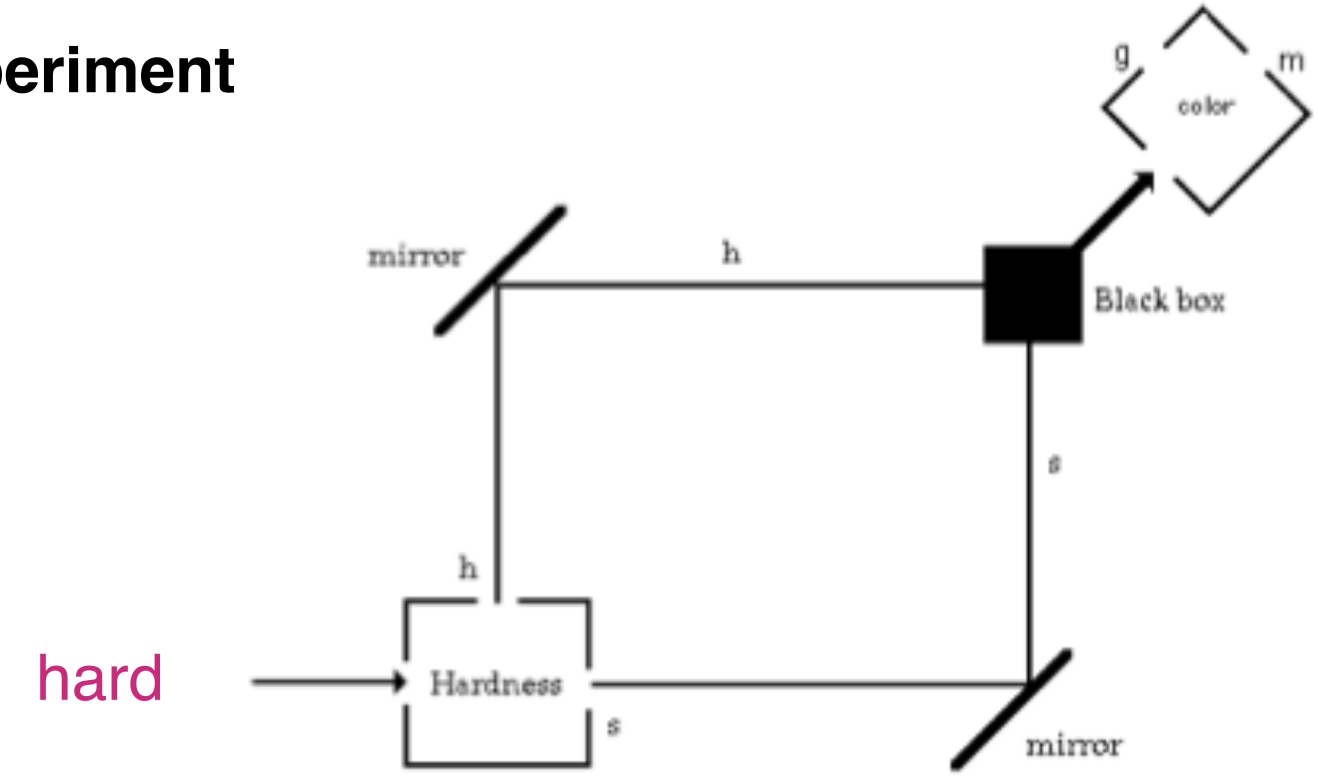
50% take h route and 50% take s route

at the end(recombined beam) 50% are hard electrons

(remember nothing bothers hardness along routes)

and 50% are soft electrons. —-> Experimentally verified!

Change experiment



Send hard electrons into hardness box.

At end add color box (**replace** final hardness box with a color box) and **measure color**.

Analysis:

All hard electrons follow h route.

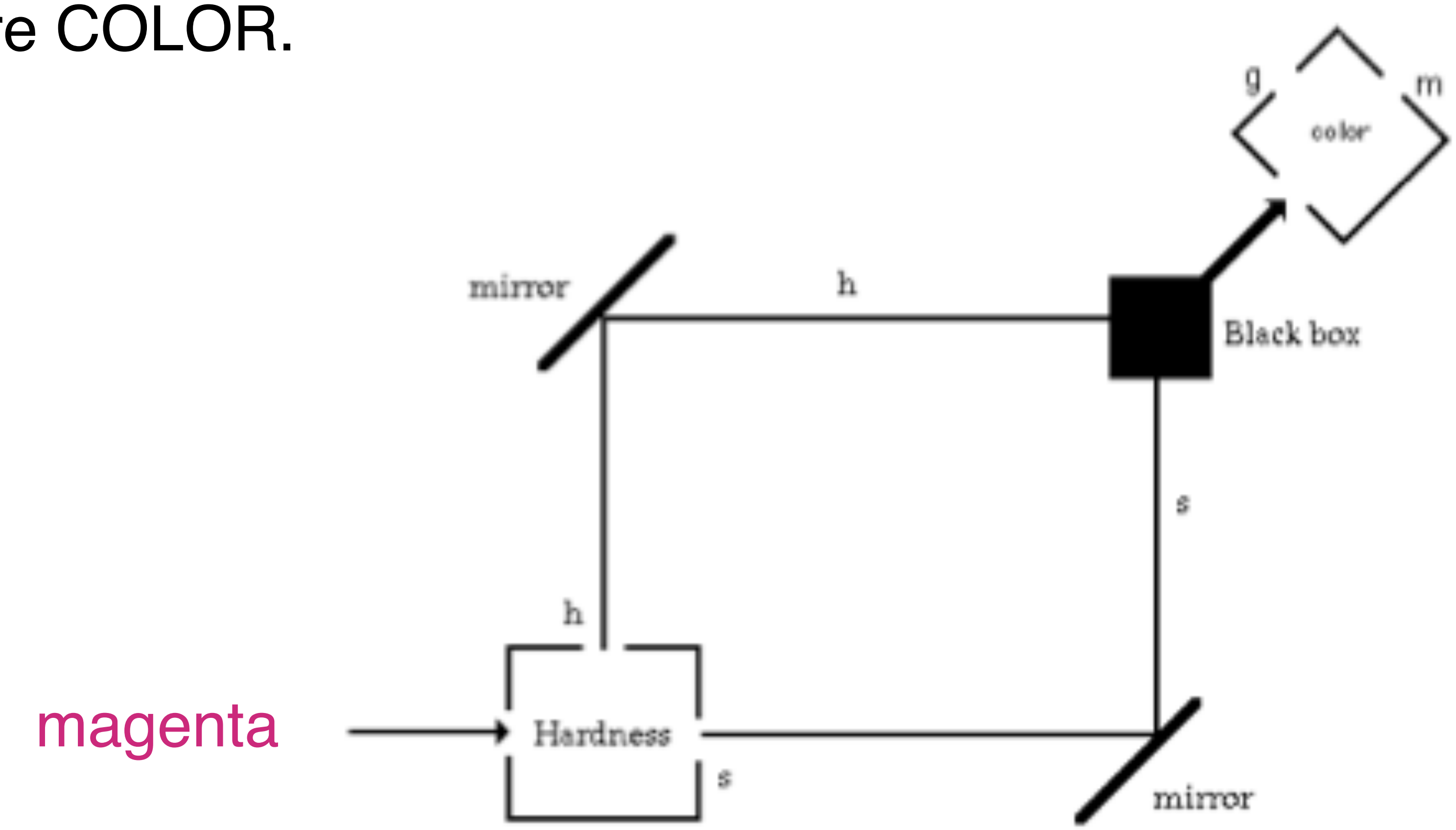
When we measure color of hard electron it is 50-50 green-magenta.

Similarly for soft electrons.

Therefore end up with 50-50 green/magenta coming out of color box.

These experimental outcomes are what would **expect** from earlier results and **do happen** exactly as described above in real world experiments. So **no** problems yet!!!

Now again send magenta electrons into hardness box, but at the end add a color box and measure COLOR.



What do you expect? The true classical physicist would give the following argument.

Since, for magenta electrons,

50% are hard and 50% are soft

(what happens to magenta electrons when sent into hardness box —> are randomized)

and each kind of electron takes the appropriate h and s routes,

at end, 50% of h (those on h route) electrons or 25% of total are magenta

and 50% of s (those on s route) electrons or 25% of total are magenta.

Thus, for 100% magenta sent in,

classical reasoning says that only 50% are magenta at end. A prediction!

Seems like a valid conclusion(prediction) since hardness boxes(1st box) **supposedly** randomize color.

Problem:

This last part of story, which a classical mind desperately wants to accept as correct, is **false**.

When we actually do experiment, **all(100%)** of electrons at end are magenta!!!

Very odd!

Hard to imagine (**classically**) what can possibly be going on in this system.

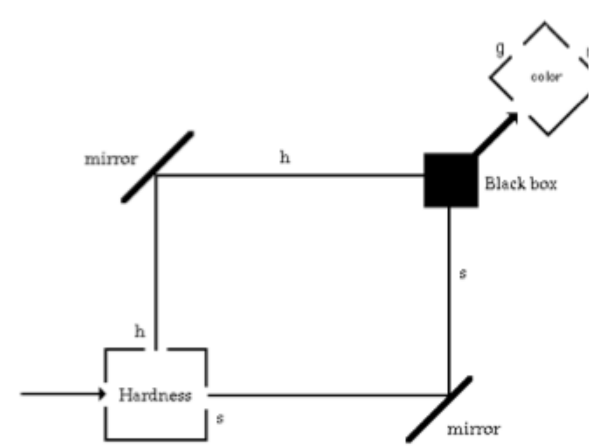
Of course, maybe it is just that our **classically oriented minds**

cannot imagine what is happening and need to be **retrained**!!!

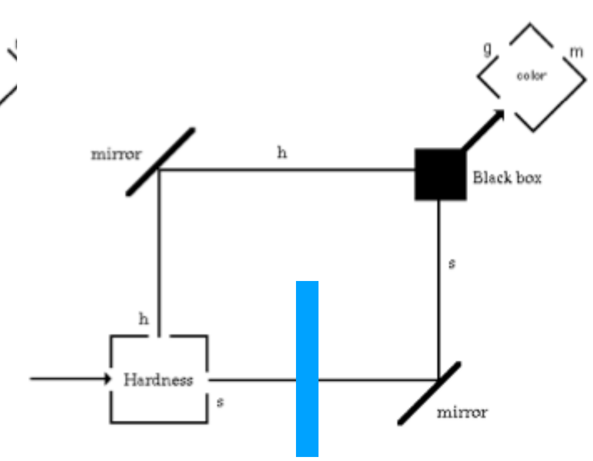
Let us **try another experiment** in hopes of understanding what is going on.

Rig up small, movable, electron-stopping wall that can be inserted in and out of route s.

When the wall **out**, we just have the earlier apparatus.



wall not inserted



wall inserted

When wall is **in**, **only** those electrons moving along h route

get through apparatus to the beam recombination device.

The electrons on the s route are blocked.

Analysis: First, there are 50% less electrons reaching beam recombination device.

When wall out all(100%) of electrons that get to beam recombination device

are magenta at end (earlier experimental result).

That means (**or seems to mean**) that all electrons that take s route

(**or we think are taking s route**) end up magenta

However, we did not check this!!!

and all that take h route (**or we think are taking h route**) end up magenta

However, we did not check this!!!

at beam recombination device when **no wall is inserted**.

This means that with wall inserted, we should end up with 50%

(1/2 lost impacting inserted wall) electrons at beam recombination device.

Based on the earlier experiment results, they should **all** be magenta

since an inserted wall on soft path **cannot possibly affect** electrons on h path

and they **all** came out magenta in the previous experiment!

This is an assumption called **locality**, which means the following:

We can separate the hard and soft paths by a very large distance

and therefore make sure that nothing we do on soft path(like inserting wall)

can affect electrons on hard path during time they are on hard path.

This is possible because the maximum speed of information transfer is

speed of light from the study of relativity.

What is actual result?

Again we are wrong in our classical way we are analyzing things. The result is **50/50 again**.

h route beam(**if there really is one**) seems to end up randomized(50-50 green-magenta)

and 50% of 50% is 25% 25 magenta and 25 green at end!

If we insert a wall in h route, **same thing happens**. Still end up only with 25% magenta!

Clearly, we have **real** trouble with our classical way of reasoning!

Seems if check (**do measurement**)

to **see which path** electrons are passing through device

(i.e., if we check to see whether the magenta electron

is passing through apparatus as a hard or a soft electron)

we get 25% magenta (50% change to green) at end

and if we DO NOT check we get 100% magenta (0% change to green).

Our classical minds are **spinning** at this point!!

We now turn to a quantum system **containing only one particle at any instant of time** instead of using beams with many particles at any instant of time.

This was not possible during the early formative day of quantum mechanics.

It only became possible in the 1980s.

Now consider a **single magenta electron** passing through apparatus when wall is out.

Can it have taken h route?

No, since all such electrons are hard and hard electrons are 50-50 green-magenta (and need 100% magenta).

If on hard path, then they are hard - that is a measurement.

Can it have taken s route?

No, since all such electrons are soft and soft electrons are 50-50 green-magenta (and need 100% magenta).

If on soft path, then are soft - it is measurement.

Can it somehow have taken **BOTH** routes at same time?

If **look** (**measure**), then half the time find electron on h route

and half the time find electron on s route,

but never find two electrons in apparatus,

or two halves of single, split electron,

one-half on each route, or anything like that.

No experimental way in which electron seems to be taking **both** routes simultaneously.

Therefore, as physicists, must **rule out** this possibility.

This eventually leads to the hypothesis for the meaning of the the word "quantum" as we will see.

Can it have taken **neither** route (got there some other way)?

Certainly not.

If we put walls in **both** routes, then NOTHING gets through at all.

Thus, the results **had** to have something to do with the box and the paths.

Summarize the dilemma for classical view.

Electrons passing through apparatus, based on measurements

do not take route h

and do not take route s

and do not take both routes at same time

and do not take neither of routes,

i.e., electrons have **zero probability** for doing any of these things (**from experiment**).

Problem is that those 4 possibilities are simply **all** of the logical possibilities

we have any notion of how to entertain using the **everyday language** of classical physics!

What can these electrons be doing?

They **must** be doing something which has simply never been thought of before in classical physics.

Electrons (in the microworld) seem to have modes of being, or modes of moving,

available to them which are unlike anything

we can think about using **words** derived from everyday ideas and classical physics.

The name of new mode (a name for something do not understand at the moment) is

SUPERPOSITION.

In the end all we can say about initially magenta electron, which is passing through our apparatus with the wall out,

Prepare yourself for this!

it is **NOT** on h path and **NOT** on s path and **NOT** on both paths and **NOT** on neither,

it is in a **SUPERPOSITION** of being on h and being on s.

What this last statement means, other than none of above, we do not know at this time;

it requires a theory to fill in the gaps and we will do that.

That is what one does in this (or any) quantum theory class.

Final and most confusing experiment.

New experimental box = **do-nothing box** - It has just 2 apertures.

An electron goes into one aperture and emerges from other aperture

with **ALL measurable properties** (color, hardness, energy, momentum, whatever)

UNCHANGED. It is Experimentally verified!

Box does not measure anything.

If we measure the properties of electrons before entering box and then after passing through box,

no measured properties are changed.

Also the time it takes for electron to get through box

is **identical** to time it would have taken if box were not there.

So nothing mysterious **seems** to be delaying it/messing around with it while in box.

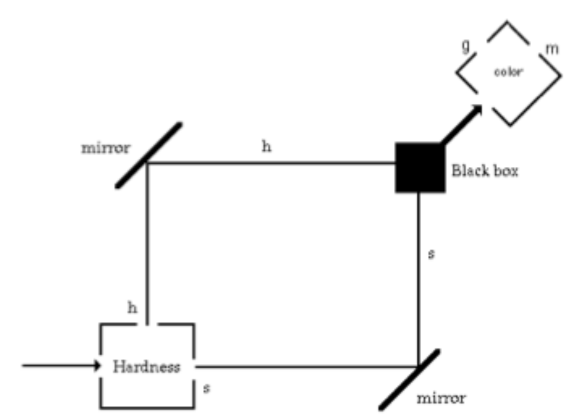
Such a **do-nothing box** can actually be constructed in the laboratory.

Experimentally if we build such a do-nothing box, then,

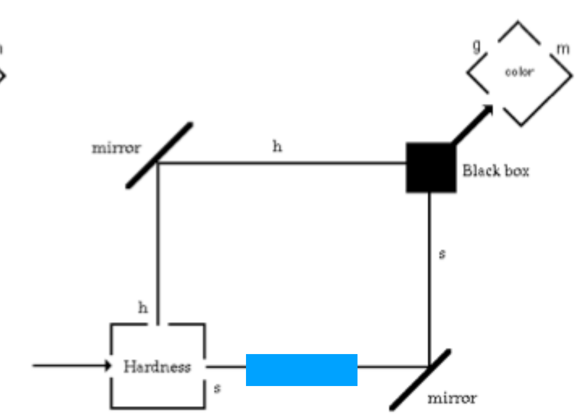
when inserted into **either** one of two paths (h or s),

it will **change all** electrons that were **magenta** at end

into **green** electrons at **end**.



box not inserted



box inserted

When the box is removed, they all go **back** to all being magenta.

So inserting box into only **one** path changes the color of electrons passing through the entire apparatus.

What is going on here?

Do-nothing boxes do not change any measurable property of electrons

that pass through it

AND of course,

do-nothing boxes do not change any measurable property of electrons

that DO NOT pass through.

That would not make any **sense** at all.

So once again our only explanation will go like......

It is not possible for the electron to have passed through the do-nothing box

since we already said that cannot change anything.

It is not possible for the electron to have passed outside the box

since the box certainly does not have a chance to bother anything

that does not even pass through it

(even though it would not do anything anyway).

It is not possible that the electron passes

both simultaneously inside and outside of the box or neither as before.

Only answer is that the electron passes through our apparatus

in a **superposition** of passing through do-nothing box

and not passing through do-nothing box

and this must **cause** the color change somehow.

This theory has got to be really neat when we finally figure out what it is we are saying about the world.

Now onto mathematics to learn/ understand the appropriate language of physics, develop a theory and then we will be able figure out what is going on.

A pattern exists……

Mathematics language <—> Relationships

Relationships <—> Correlations

Correlations <—> Quantum Mechanics

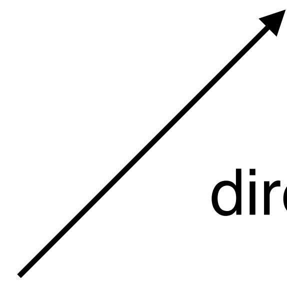
Quantum Mechanics <—> Mathematics language

The reality of this pattern will become clear as we proceed through the class.

The Mathematics Of Quantum Mechanics - Part 1

We must learn some mathematics BECAUSE it is language of QM.
We will not be able to explain and use the ideas of QM using WORDS!!

First, some High School level Stuff…

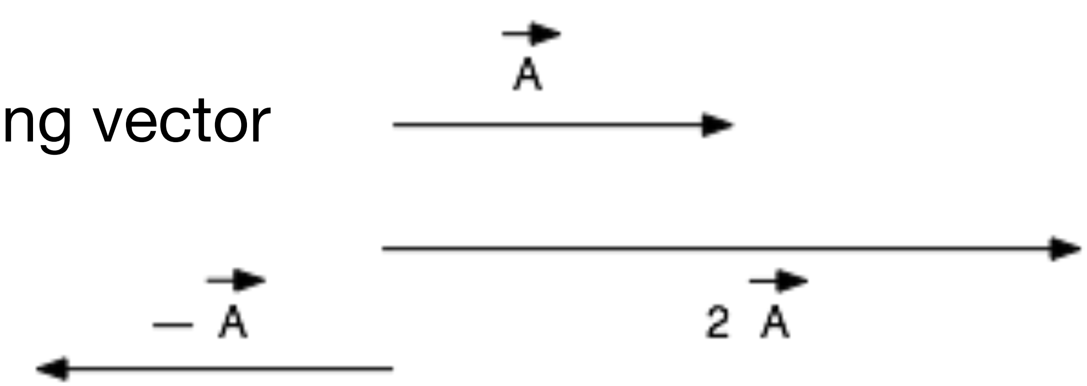
Vector =  directed line segment

if in a plane—-> 2 numbers (a,b) = number of dimensions

or Vector —> 3 numbers (a,b,c)= number of dimensions in ordinary space

Multiplication by a scalar(a number)

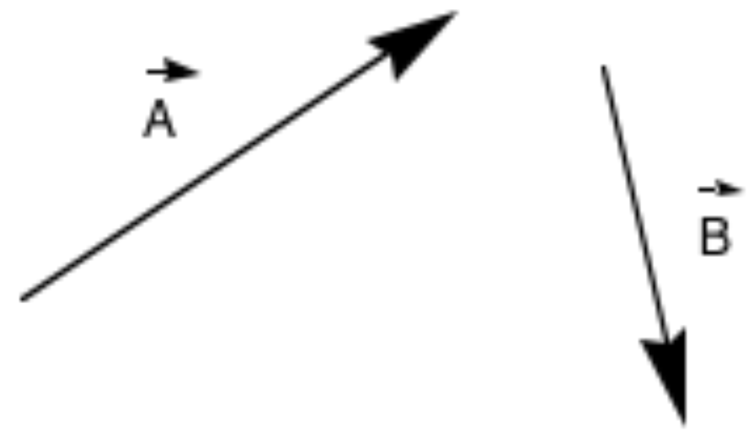
starting vector



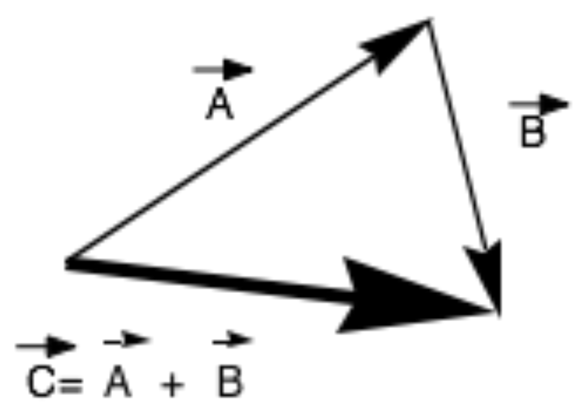
multiply by 2

multiply by -1

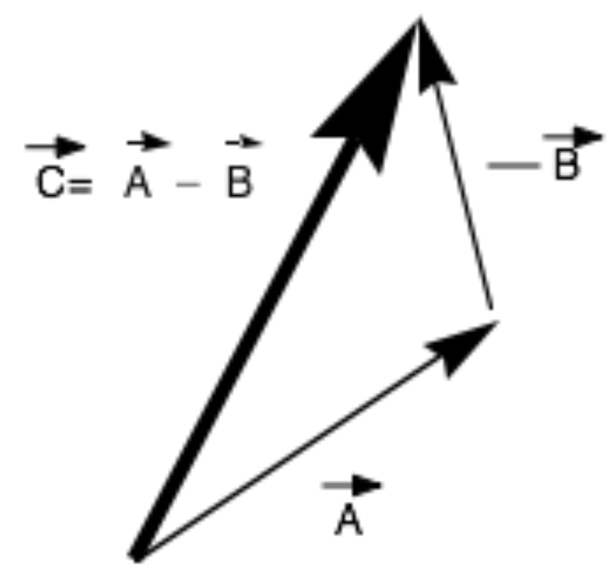
Given two vectors



one can add/subtract them with the geometric rules shown below



Sum

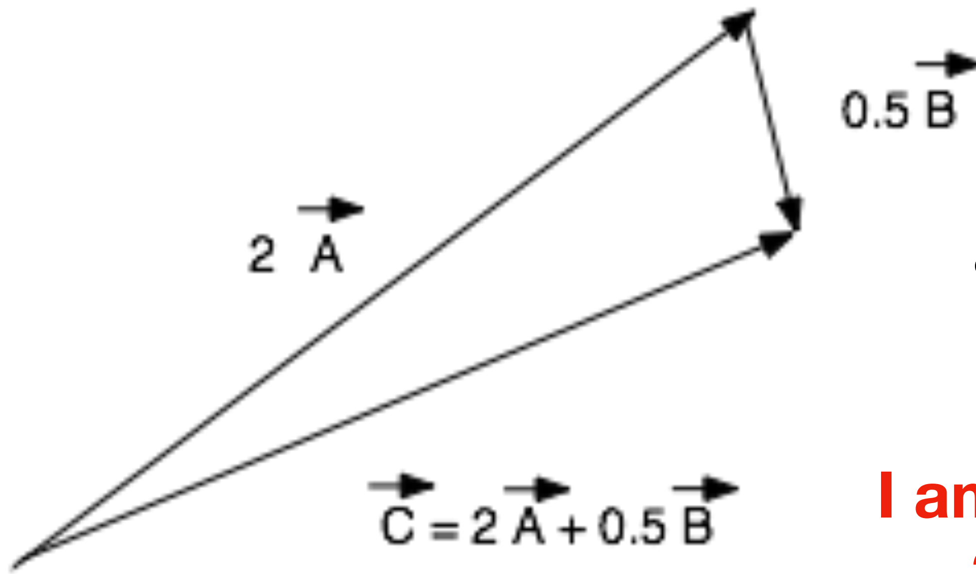


Difference

call the “tail to tip” rule

Now we generalize —> extend idea to a general linear combination(sum) of vectors

$$\vec{C} = \alpha\vec{A} + \beta\vec{B}$$ it is still a vector!



adding parts => total

using $\alpha = 2$, $\beta = 0.5$

I am drawing pictures in the real 3-dimensional world here!

Here, I note that all coefficients α, β are real numbers at this time.

These ideas and methods work perfectly for all macroworld vectors(high school math).

But we must generalize further in order to deal with microworld vectors(this class).

Let us see how……….

The Space of Physics

vector = 3 numbers (3 = dimension of our world)

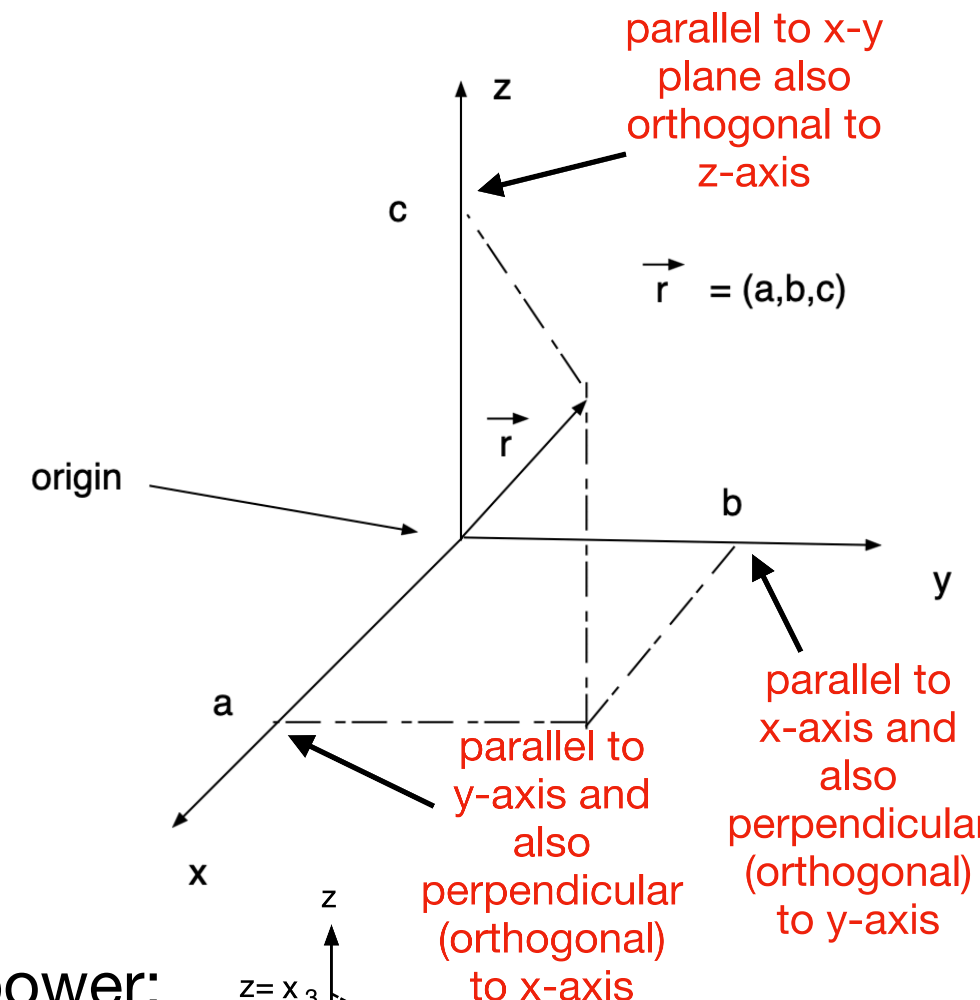
Location in space (relative to an origin) = 3 numbers

use axis labels $(x = x_1, y = y_1, z = z_1)$

—-> can define a position or radius vector

$$\vec{r} = (x, y, z) = (x_1, x_2, x_3) \; = (a, b, c)$$

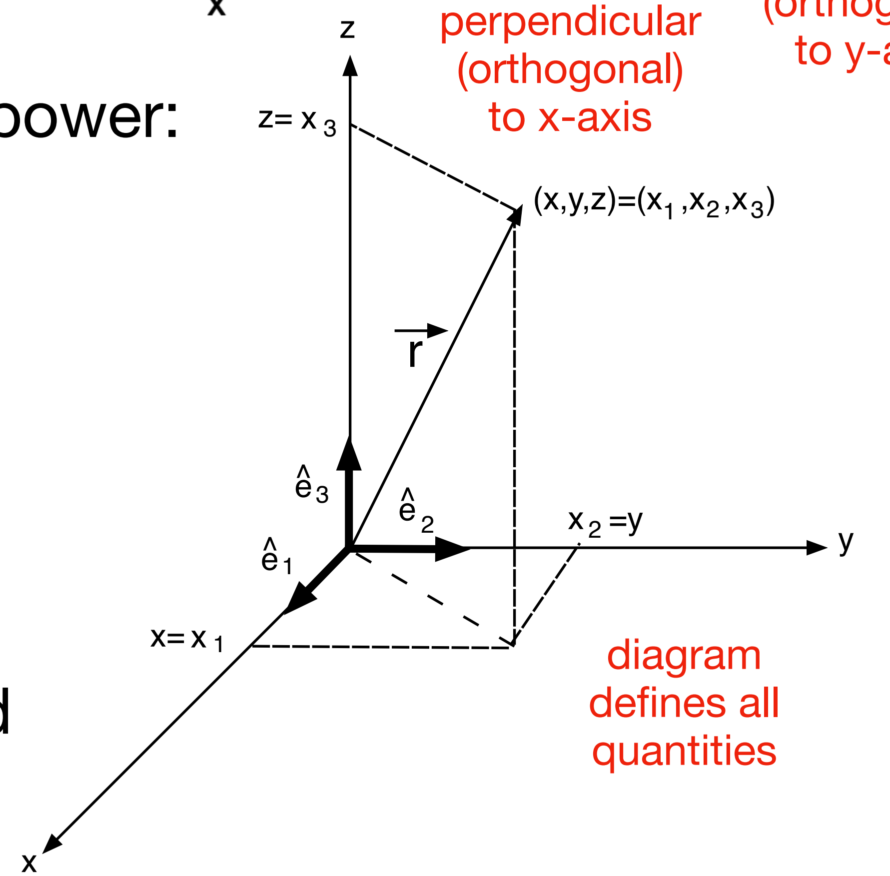
as shown in diagram



Now to add more mathematics to increase the descriptive power:

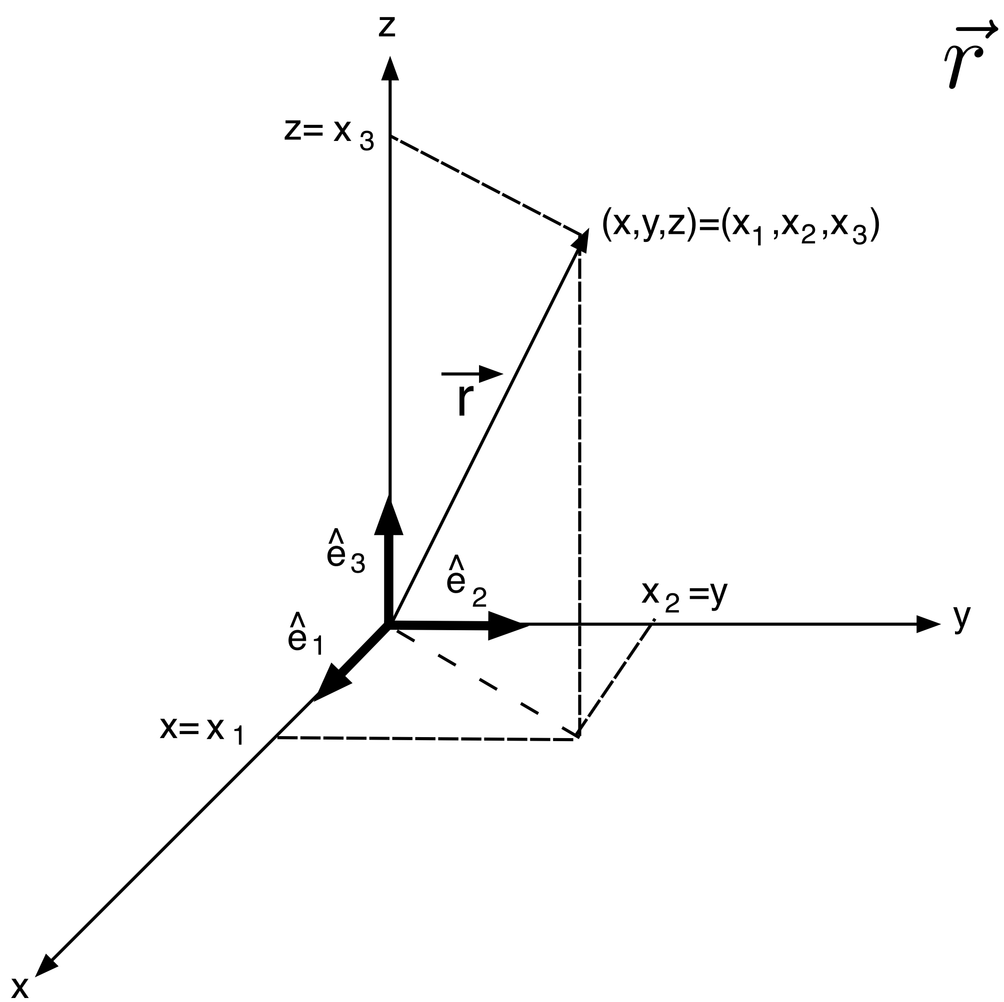
First - define 3 special vectors (called a **basis**) that will be used to construct all other vectors - several different label sets….

$$\hat{e}_1, \hat{e}_2, \hat{e}_3 \text{ or } \hat{i}, \hat{j}, \hat{k} \text{ or } \hat{x}, \hat{y}, \hat{z}$$



They are unit vectors (length=1) and define a Right-Handed coordinate system as shown.

By using vector addition rules we have (could prove this using tip-to-tail rules)



$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k} = x_1\hat{e}_1 + x_2\hat{e}_2 + x_3\hat{e}_3$$
$$= x\hat{x} + y\hat{y} + z\hat{z}$$

In words, x_1=x=component of $\vec{r}$ in 1-direction(x-direction), etc

$x_1\hat{e}_1$ **is a vector in 1-direction of length x_1**

$\vec{r}$ **is vector sum of three vectors** $x_1\hat{e}_1$ **,** $x_2\hat{e}_2$ **,** $x_3\hat{e}_3$ **draw it yourself**

$\hat{e}_1, \hat{e}_2, \hat{e}_3$ or $\hat{i}, \hat{j}, \hat{k}$ or $\hat{x}, \hat{y}, \hat{z}$ **called a basis set**

Important fact —> any vector can be written in terms of a basis

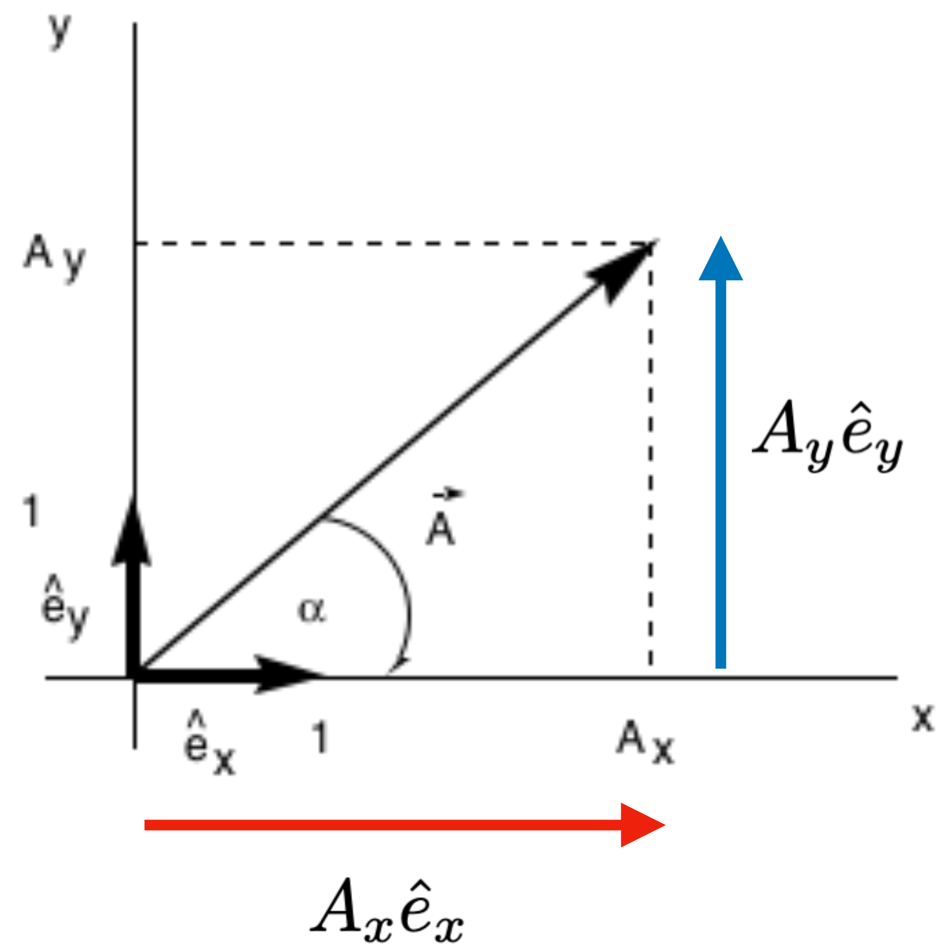
vector length $r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2}$

using Pythagorean theorem

Go to 2 dimensions (for drawing **simplicity)**
Everything I do generalizes to higher dimensions easily.

$$\hat{e}_x = \text{ unit(length } = 1) \text{ vector in } x - \text{direction}$$

$$\hat{e}_y = \text{ unit(length } = 1) \text{ vector in } y - \text{direction}$$

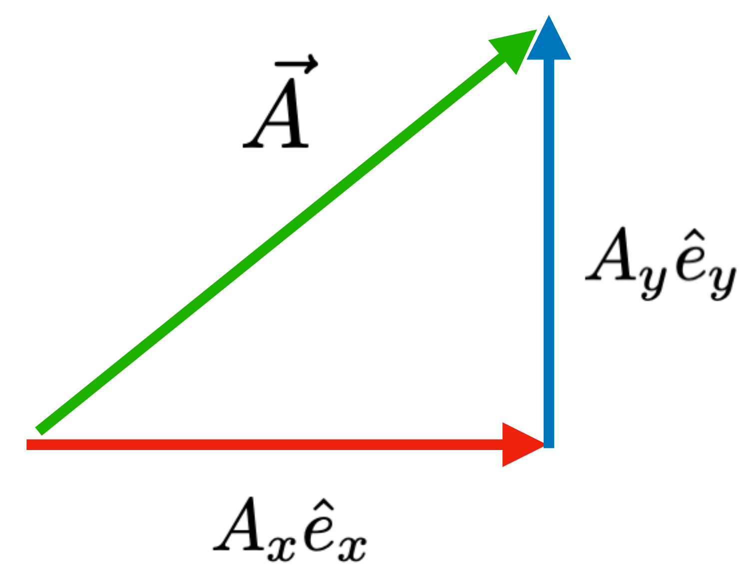


multiply by a scalar

$$A_x \hat{e}_x = \text{ vector of length } A_x \text{ in the } x - \text{direction}$$

$$A_y \hat{e}_y = \text{ vector of length } A_y \text{ in the } y - \text{direction}$$

$$\vec{A} = A_x \hat{e}_x + A_y \hat{e}_y$$

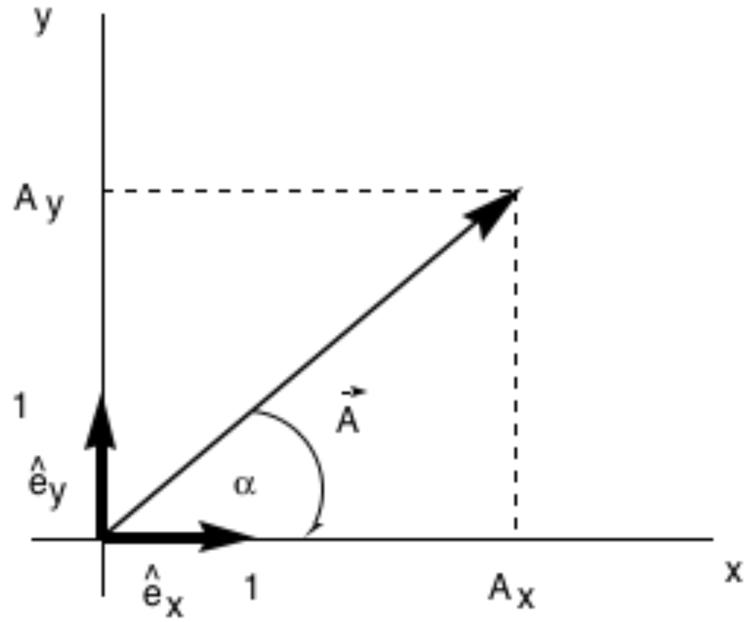


by addition rule for vectors as shown

Now from the diagram

$$A_x = A\cos\alpha \quad , \quad A_y = A\sin\alpha$$

trigonometry —> **components**



For two vectors we now have

$$\vec{A} = A_x\hat{e}_x + A_y\hat{e}_y \quad , \quad \vec{B} = B_x\hat{e}_x + B_y\hat{e}_y$$

$$\vec{A} + \vec{B} = (A_x + B_x)\,\hat{e}_x + (A_y + B_y)\,\hat{e}_y$$

$$\vec{A} - \vec{B} = (A_x - B_x)\,\hat{e}_x + (A_y - B_y)\,\hat{e}_y$$

thus the old tip-to-tail idea is now **eliminated** in favor of adding and subtracting components; tip-to-tail idea cannot be extended to more complicated situations such as in microword but the new idea **can**!!

More Generalization - Further thinking about vectors

old $\quad \vec{V} = v_1\hat{e}_1 + v_2\hat{e}_2 \quad$ or $\quad |V\rangle = (v_1, v_2)$

just a change of "notation" —> removes unit vectors

$\longrightarrow \quad |V\rangle = (v_1, v_2) \quad \longrightarrow \quad |A\rangle = (A_x, A_y)$

Start of Dirac language

new symbol $|\ \rangle$ —> Dirac **Ket** vector - it is the **first** element in a new language

Your only job is to learn the new terminology(language); I will do all the algebraic manipulations in class.

Now using new language to replace old equations……

addition

If $|V_1\rangle = (7, -2)$ and $|V_2\rangle = (-5, 3)$

then the sum $|V\rangle = |V_1\rangle + |V_2\rangle = (7 + (-5), (-2) + 3) = (2, 1)$

we are still adding components in same way as before!

difference

If $|V_1\rangle = (7, -2)$ and $|V_2\rangle = (-5, 3)$

then the difference $|V\rangle = |V_1\rangle - |V_2\rangle = (7 - (-5), (-2) - 3) = (12, -5)$

we are still subtracting components in same way as before!

length

If $|V\rangle = (v_1, v_2)$ then length $V = \sqrt{v_1^2 + v_2^2}$

Now we have the language elements that will allow us to extend ideas so they are useful in the microword

We now define a **new operation** and the **generalized vector length**

$$\text{If } \; |V\rangle = (v_1, v_2) \quad , \quad |U\rangle = (u_1, u_2) \,,$$

$$\text{then } \; \langle V \,|\, U\rangle = v_1 u_1 + v_2 u_2$$

definition

Here we are using vectors with real-valued components —> called **Euclidean vectors**

the symbol $\langle \cdots | \cdots \rangle$

—> Dirac **bra**-c-**ket** or **braket** symbol —> another element in Dirac language

Then we have

$$\langle V \,|\, V\rangle = v_1 v_1 + v_2 v_2 = v_1^2 + v_2^2 = (\text{length})^2$$

$$\text{length } = \sqrt{\langle V \,|\, V\rangle} = \sqrt{v_1^2 + v_2^2}$$

In the world we are in at this moment with real-valued component Euclidean vectors all vector lengths are non-negative numbers, i.e.,

$$\langle V \,|\, V\rangle \geq 0$$

One of the reasons for developing these new ideas is that the formalism now easily extends to more dimensions……

3 dimensions $|V\rangle = (v_1, v_2, v_3) \quad , \quad |U\rangle = (u_1, u_2, u_3)$

$$\langle V \,|\, V\rangle = v_1 v_1 + v_2 v_2 + v_3 v_3 \qquad \langle U \,|\, U\rangle = u_1 u_1 + u_2 u_2 + u_3 u_3$$

$$\langle V \,|\, U\rangle = v_1 u_1 + v_2 u_2 + v_3 u_3$$

Definition:

summation symbol $\sum_{i=n_1}^{n_2}$

$$\sum_{j=1}^{5} A_j = A_1 + A_2 + A_3 + A_4 + A_5$$

$$\sum_{i=3}^{7} A_i B^i = A_3 B^3 + A_4 B^4 + A_5 B^5 + A_6 B^6 + A_7 B^7$$

and so on ……

then we can write $\langle V \,|\, U\rangle = v_1 u_1 + v_2 u_2 + v_3 u_3 = \sum_{k=1}^{3} v_k u_k$

$$\text{length} = \sqrt{\langle V \,|\, V\rangle} = \sqrt{v_1 v_1 + v_2 v_2 + v_3 v_3} = \sqrt{\sum_{k=1}^{3} v_k^2}$$

More than 3 dimensions (we are not in Kansas anymore, Toto)

I can no longer easily visualize(draw on paper) so I stop trying!!

$$|V\rangle = (v_1, v_2, ..., v_n) \quad , \quad |U\rangle = (u_1, u_2, ..., u_n)$$

$$\langle V \mid U\rangle = v_1 u_1 + v_2 u_2 + \cdots + v_n u_n = \sum_{k=1}^{n} v_k u_k$$

extended definition

$$\text{length} \;= \sqrt{\langle V \mid V\rangle} = \sqrt{v_1 v_1 + v_2 v_2 + \cdots + v_n v_n} = \sqrt{\sum_{k=1}^{n} v_k^2}$$

Dirac Braket = scalar or inner product(or “dot” product)

Now the next step is to generalize the idea of "components"

Up to now "components" = "real" numbers and vectors were "Euclidean vectors"

Generalization —-> complex components i.e., components are complex numbers

Digression on simple properties of complex numbers

$i^2 = -1 \;,\; i^3 = -i \;,\; i^4 = +1 \;,\; i^5 = i \;,\; i^6 = -1 \;,\;$ and so on — **definition** of i

$z = a + bi$ — **definition** of complex number — a = **real** part — b= **imaginary** part

$$(7+4i)+(-2+9i) = 5+13i$$

definition of addition(or subtraction) — add real parts and imaginary parts separately

$$(7+4i)(-2+9i) = (7)(-2)+(7)(9i)+(4i)(-2)+(4i)(9i)$$
$$= -14+63i-8i-36 = -50+55i$$

using $i^2 = -1$

definition of multiplication

definition complex conjugate $z^* = a - bi$

definition absolute value

$$|z|^2 = z^* z = (a-bi)(a+bi) = a^2+b^2 \qquad |z| = \sqrt{a^2+b^2}$$

Note that if z is real(imaginary part = 0) then z* = z

In this generalized world the **braket** also has an extended definition

$$\langle V \mid U \rangle = v_1^* u_1 + v_2^* u_2 + \cdots + v_n^* v_n = \sum_{k=1}^{n} v_k^* u_k$$

If components are real(Euclidean vectors), the $v_i^*=v_i$ and we have the same formula as earlier

$$\langle V \mid U \rangle = v_1 u_1 + v_2 u_2 + \cdots + v_n u_n = \sum_{k=1}^{n} v_k u_k$$

now $\quad \text{length} \; = \sqrt{\langle V \mid V \rangle} = \sqrt{\sum_{k=1}^{n} v_k^* v_k} = \sqrt{\sum_{k=1}^{n} |v_k|^2}$

Vectors in QM = Euclidean vectors with complex components - they live in a mathematical place called a "Hilbert space"

Now to extend all these idea to include Basis Vectors

just 2 dimensions for simplicity

$$|1\rangle = (1,0) \leftrightarrow \hat{e}_x \quad , \quad |2\rangle = (0,1) \leftrightarrow \hat{e}_y$$

new representation for old unit basis vectors

$$\hat{e}_1 = \hat{e}_x \;\; \hat{e}_2 = \hat{e}_y \;\; \hat{e}_3 = \hat{e}_z$$

3 dimensions

$$|1\rangle = (1,0,0) \quad , \quad |2\rangle = (0,1,0) \quad , \quad |3\rangle = (0,0,1)$$

called n-tuples

Clearly length = 1 (unit vectors) and they are perpendicular or orthogonal(by definition) and as we will prove shortly

All the same……with changed mathematics language

Some examples:

$$|V\rangle = v_1\hat{e}_1 + v_2\hat{e}_2 + v_3\hat{e}_3$$

$$|V\rangle = (v_1, v_2, v_3)$$

$$|V\rangle = v_1\,|1\rangle + v_2\,|2\rangle + v_3\,|3\rangle$$

$$|1\rangle = (1, 0, 0)$$

$$\langle 1\,|\,V\rangle = v_1\,\langle 1\,|\,1\rangle + v_2\,\langle 1\,|\,2\rangle + v_3\,\langle 1\,|\,3\rangle$$

$$= v_1(1) + v_2(0) + v_3(0) = v_1$$

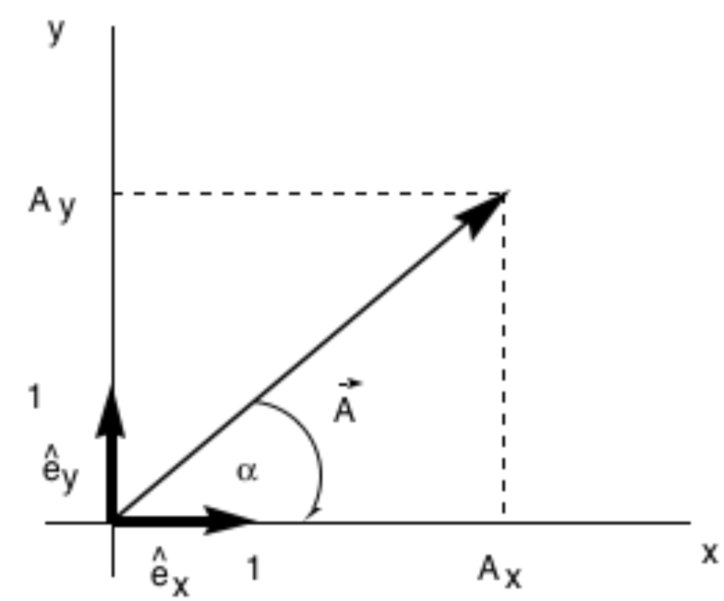
Definition: Vector Space = collection of vectors such that

if add any two vectors together => another vector in collection;

and have **a scalar product(bracket)** defined

—> Hilbert space in microworld

Can write **any** vector in terms of **basis <—> orthonormal set**, i.e.,



no longer a valid picture when components are complex; just suggestive!!

$$|v_1, v_2\rangle = v_1(1,0) + v_2(0,1) = (v_1, 0) + (0, v_2) = (v_1, v_2)$$

$$|v_1, v_2\rangle = v_1(1,0) + v_2(0,1) = v_1 |1\rangle + v_2 |2\rangle$$

$$\vec{v} = (v_1, v_2) = v_1\hat{e}_1 + v_2\hat{e}_2$$ **“old way”**

v_1 = component of vector in 1-direction, etc

Similarly for 3 dimensions

$$|U\rangle = (u_1, u_2, u_3) = u_1(1,0,0) + u_2(0,1,0) + u_3(0,0,1) = u_1 |1\rangle + u_2 |2\rangle + u_3 |3\rangle$$

and so on

Using earlier rule for evaluating the scalar product (sum of component products)

$$\langle 1 \,|\, 1 \rangle = (1,0)\cdot(1,0) = (1)(1) + (0)(0) = 1 \qquad \langle 1 \,|\, 2 \rangle = (1,0)\cdot(0,1) = (1)(0) + (0)(1) = 0$$

$$\langle 2 \,|\, 1 \rangle = (0,1)\cdot(1,0) = (0)(1) + (1)(0) = 0 \qquad \langle 2 \,|\, 2 \rangle = (0,1)\cdot(0,1) = (0)(0) + (1)(1) = 1$$

$$|V\rangle = v_1\,|1\rangle + v_2\,|2\rangle$$

=1 **=0** **=0** **=1**

$$\langle 1 \,|\, V \rangle = v_1\,\langle 1 \,|\, 1 \rangle + v_2\,\langle 1 \,|\, 2 \rangle \qquad \langle 2 \,|\, V \rangle = v_1\,\langle 2 \,|\, 1 \rangle + v_2\,\langle 2 \,|\, 2 \rangle$$

$$\longrightarrow \quad \langle 1 \,|\, V \rangle = v_1 \qquad \langle 2 \,|\, V \rangle = v_2$$

in new language

—> The component of vector in particular direction(along a basis vector) is given by the scalar product(bracket) of vector with corresponding basis vector. New Definition of component!

Remember component = projection on the axis in old real world!

But now, in Hilbert space, it is a "braket" - no more pictures!!

Thus, we can always write

$$|V\rangle = \underbrace{\langle 1 \,|\, V \rangle}_{v_1}\,|1\rangle + \underbrace{\langle 2 \,|\, V \rangle}_{v_2}\,|2\rangle$$

general definition in new language

Orthogonality

First way - remember High School - using old language

$$\vec{V} = (7,4) \qquad \hat{x} = (1,0) \quad , \quad \hat{y} = (0,1)$$

$$\longrightarrow \vec{V} = 7\hat{x} + 4\hat{y} \qquad \textbf{using old notation}$$

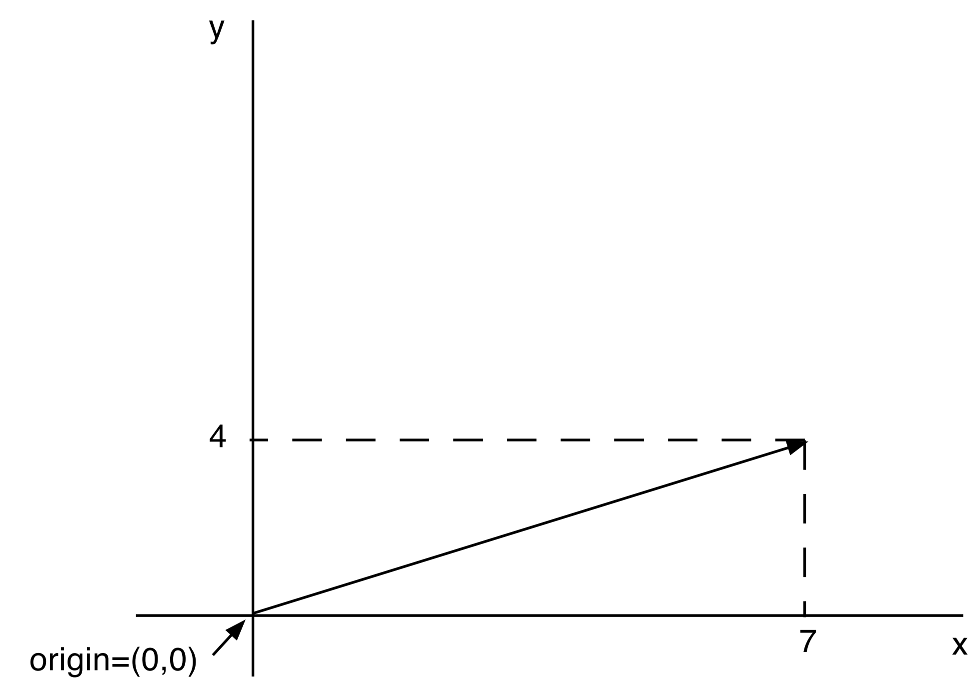
Define old “dot” product:

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y \; = \langle A \,|\, B \rangle$$

components

$$7 = v_x = \hat{x} \cdot \vec{V} \quad , \quad 4 = v_y = \hat{y} \cdot \vec{V}$$

$$\sqrt{v_x^2 + v_y^2} = \sqrt{65} \quad \text{length}$$



Notice that every equation is identical to those in Dirac language. We only went to new language because these older ideas cannot be extended into microworld!

Now we can always write (in peculiar way - kind of stuff mathematicians and theoretical physicists do all the time!)

$$7 = v_x = \hat{x}\cdot\vec{V} = length(\hat{x}) \times length(\vec{V}) \times \frac{v_x}{length(\vec{V})}$$

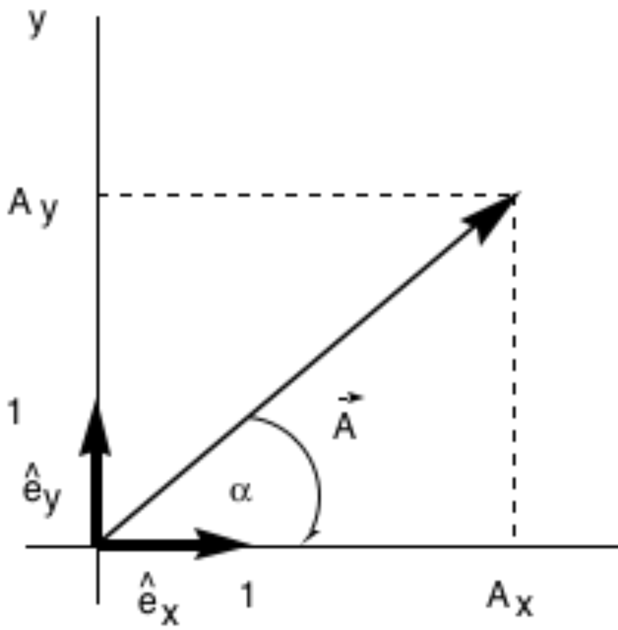
whole mess just multiplies v_x by 1

Can you see?

length =1

What are you really doing when you carry out algebra? You are either multiplying by 1 or adding 0 — otherwise you will change the expression you are working on!

Now $$\frac{v_x}{length(\vec{V})} = \text{cosine of angle between } \vec{V} \text{ and } \hat{x}$$ **—> definition of cosine**



so $$\vec{A}\cdot\vec{B} = length(\vec{A}) \times length(\vec{B}) \times \cos\left(angle\, between\, \vec{A}\, and\, \vec{B}\right)$$

generalize ⟶ $$\vec{A}\cdot\vec{B} = AB\cos\theta$$ ⟶ **0 if perpendicular**

Important special cases:

$$\hat{x}\cdot\hat{x} = length(\hat{x}) \times length(\hat{x})\cos\left(0^\circ\right) = 1 \quad \hat{x}\cdot\hat{y} = length(\hat{x}) \times length(\hat{y})\cos\left(90^\circ\right) = 0$$

—> Basis vectors are orthonormal! As I have been assuming!

= orthogonal + normalized to 1(length=1) —> orthonormal

Generalize further (add new notation) by defining in a vector space

$$\langle A \,|\, B\rangle = AB\cos\left(\theta_{AB}\right)$$

orthonormality clear!

no simple geometric picture possible (as with real components); "angle" is not geometric angle anymore, but just a some kind of parameter now!

To easily explore further we need the **definition** of a new object - the **Kronecker Delta.**

$$\langle n \,|\, m\rangle = \delta_{nm} = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}$$

for basis vectors $|n\rangle\,,|m\rangle$

definition just says that basis vector are unit length + orthogonal or orthonormal

then this implies that

$$\langle 1 \,|\, 1\rangle = \delta_{11} = 1$$ **—-> length =1**

$$\langle 1 \,|\, 2\rangle = \delta_{12} = 0$$ **—-> orthogonal**

$$\langle 2 \,|\, 1\rangle = \delta_{21} = 0$$ **—-> orthogonal**

$$\langle 2 \,|\, 2\rangle = \delta_{22} = 1$$ **—-> length =1**

Using Kronecker delta to do algebra -> very powerful tool that I will use to do manipulations and you will use if you make the effort to learn how to do the algebra.

Examples

$$\sum_{k=1}^{3} A_k \delta_{k2} = A_1\delta_{12} + A_2\delta_{22} + A_3\delta_{32} = A_1(0) + A_2(1) + A_3(0) = A_2$$

$$\sum_{k=1}^{n} A_k \delta_{km} = A_m \quad m \le n$$

so appearance of kronecker delta inside a summation eliminates one summation —> a very powerful algebraic tool !!

In order to complete the mathematical tools needed to describe the microworld we need one last mathematical object —> completely new for some of you guys —> a **Matrix.**

Definition

m × n matrix is **m × n array (m rows and n columns)** of numbers(**matrix elements**) with a well-defined set of associated mathematical rules.

For example, see sample matrices shown below

$$\begin{pmatrix} 2 \\ 5 \end{pmatrix}$$ 2 elements

2 × 1 matrix (= column vector)

$$\begin{pmatrix} 2 & 5 \end{pmatrix}$$ 2 elements

1 x 2 matrix (= row vector)

$$\begin{pmatrix} 2 & 5 \\ -3 & 10 \end{pmatrix}$$ 4 elements

2 × 2 matrix

$$\begin{pmatrix} 2 & 5 & -1 \\ -3 & 10 & 5 \end{pmatrix}$$

2 × 3 matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad , \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

now using row-column labels

diagonal elements   off-diagonal elements

Consider two matrices

$$\mathbf{A} = \begin{pmatrix} 2 & 5 \\ -3 & 10 \end{pmatrix} \quad , \quad \mathbf{B} = \begin{pmatrix} -8 & 15 \\ 7 & 4 \end{pmatrix}$$

—> two 2 × 2 matrices (note **boldface**)

addition => add elements

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} (2) + (-8) & (5) + (15) \\ (-3) + (7) & (10) + (4) \end{pmatrix} = \begin{pmatrix} -6 & 20 \\ 4 & 14 \end{pmatrix}$$

matrix multiplication

$$\mathbf{AB} = \begin{pmatrix} 2 & 5 \\ -3 & 10 \end{pmatrix} \begin{pmatrix} -8 & 15 \\ 7 & 4 \end{pmatrix} = \begin{pmatrix} (2)(-8) + (5)(7) & (2)(15) + (5)(4) \\ (-3)(-8) + (10)(7) & (-3)(15) + (10)(4) \end{pmatrix} = \begin{pmatrix} 19 & 50 \\ 94 & -5 \end{pmatrix}$$

I just followed the **general rule**

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad , \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

using row-column labels

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

These following are definitions so that you know what I mean when I refer to some matrix property.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Special cases:

definition

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = identity\ matrix$$ **–>** $$\mathbf{IA} = \mathbf{AI} = \mathbf{A}$$ effect = no change

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21} = determinant\, of\ \mathbf{A}$$

definitions only for 2 x 2 matrices

definition

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = inverse\ matrix$$ **–>** $$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

effect

Example:

$$A = \begin{pmatrix} 1 & 4i \\ -4i & 7 \end{pmatrix} \qquad A^{-1} = \frac{1}{2 - 12i} \begin{pmatrix} 2 & -4i \\ -3 & 1 \end{pmatrix}$$ **–>** $$AA^{-1} = I$$

$$\mathbf{A}^{\mathbf{T}} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} = transpose\ matrix \qquad \mathbf{A}^{\dagger} = \begin{pmatrix} a_{11}^{*} & a_{21}^{*} \\ a_{12}^{*} & a_{22}^{*} \end{pmatrix} = Hermitian\ conjugate\ matrix$$

definition definition

If $\mathbf{A}^{\dagger} = \mathbf{A}$ matrix = Hermitian

$$A = \begin{pmatrix} 1 & 4i \\ -4i & 7 \end{pmatrix}$$

If $\mathbf{A}^{\dagger} = \mathbf{A}^{-1}$ matrix = Unitary

$$A = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}$$

Finally $$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$ **–>** $$Tr(\mathbf{A}) = a_{11} + a_{22}$$ **= sum over diagonal elements**

Trace

Now start to pull it all together for QM.

or we now broaden the Dirac language for use in QM

lots of jokes…
Dirac was oblivious!

$$|V\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \text{“}ket'' \; vector$$

column vector

$$\langle V| = \begin{pmatrix} v_1^* & v_2^* \end{pmatrix} = \text{“}bra'' \; vector$$

new definition

row vector

Every Ket vector has a corresponding Bra vector in the Hilbert space.

Then real mathematical meaning of the “braket” is

$$\langle V \,|\, U\rangle = \begin{pmatrix} v_1^* & v_2^* \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = v_1^* u_1 + v_2^* u_2 = \text{“}braket''$$

—> rule is just matrix multiplication

From now on we assume that all ket/bra vectors are unit vectors(required for QM to be valid in the microworld as we will see)

Finally we need to think about Operators

An **operator** is a mathematical object which acts on any vector in our vector space and **results in another vector in same vector space**.

Its operation and behavior is similar to a **function with numbers** y = f(x)

insert number in function —> new number

Example

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad , \quad |V\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$\mathbf{A}\,|V\rangle = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = |U\rangle$$

just another vector in the space!!

Therefore

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$ = an operator **—> matrices are operators in QM**

why I needed to introduce the idea!

Operators are very important in QM because

observables or measurable quantities are represented in QM by Hermitian operators(matrices) where $\mathbf{A}^{\dagger} = \mathbf{A}$

and

transformations or physical changes of vectors are represented in QM by Unitary operators (matrices) where $\mathbf{A}^{\dagger} = \mathbf{A}^{-1}$

Properties of operators - $\hat{\mathbf{O}}$ (note boldface + hat)

Matrix representation (I show you this so you know what it is when I refer to it)

you will not need to do such manipulations yourself!!

Definition: Matrix element

$o_{nm} = \langle n | \hat{\mathbf{O}} | m \rangle$ = (nm) matrix element of $\hat{\mathbf{O}}$ (nm) = (row,column)

—-> matrix representing $\hat{\mathbf{O}}$ is

Using the $|1\rangle, |2\rangle$ basis

$$\mathbf{O} = \begin{pmatrix} \langle 1 | \hat{\mathbf{O}} | 1 \rangle & \langle 1 | \hat{\mathbf{O}} | 2 \rangle \\ \langle 2 | \hat{\mathbf{O}} | 1 \rangle & \langle 2 | \hat{\mathbf{O}} | 2 \rangle \end{pmatrix} = \begin{pmatrix} o_{11} & o_{12} \\ o_{21} & o_{22} \end{pmatrix}$$

Now assume that $\hat{\mathbf{O}} |1\rangle = |q\rangle$, $\hat{\mathbf{O}} |2\rangle = |r\rangle$ that is what operators do!

then

$$o_{11} = \langle 1 | \hat{\mathbf{O}} | 1 \rangle = \langle 1 | q \rangle$$

$$o_{12} = \langle 1 | \hat{\mathbf{O}} | 2 \rangle = \langle 1 | r \rangle$$

$$o_{21} = \langle 2 | \hat{\mathbf{O}} | 1 \rangle = \langle 2 | q \rangle$$

$$o_{22} = \langle 2 | \hat{\mathbf{O}} | 2 \rangle = \langle 2 | r \rangle$$

—-> Matrix elements are just brackets = numbers!!

Most important special case:

If $\hat{\mathbf{A}}\,|V\rangle = a\,|V\rangle$ (i.e., we get same vector back) then

$|V\rangle$ called an **eigenvector** of operator $\hat{\mathbf{A}}$ with **eigenvalue** a .

For Hermitian operators, used for observables in QM, the set of eigenvectors **always** forms an orthonormal basis -> called a **complete set** and the eigenvalues are all **real**. Very useful!

If basis set that is used to calculate matrix representing an operator

is the set of eigenvectors of the operator, i.e., if

$$\hat{\mathbf{O}}\,|1\rangle = o_1\,|1\rangle \quad , \quad \hat{\mathbf{O}}\,|2\rangle = o_2\,|2\rangle \qquad \{o_1, o_2\} \quad \text{= eigenvalues}$$

then

$$o_{11} = \langle 1|\,\hat{\mathbf{O}}\,|1\rangle = o_1\,\langle 1\,|\,1\rangle = o_1 \quad o_{21} = \langle 2|\,\hat{\mathbf{O}}\,|1\rangle = o_1\,\langle 2\,|\,1\rangle = 0$$

$$o_{12} = \langle 1|\,\hat{\mathbf{O}}\,|2\rangle = o_2\,\langle 1\,|\,2\rangle = 0 \quad o_{22} = \langle 2|\,\hat{\mathbf{O}}\,|2\rangle = o_2\,\langle 2\,|\,2\rangle = o_2$$

$$\longrightarrow \mathbf{O} = \begin{pmatrix} \langle 1|\,\hat{\mathbf{O}}\,|1\rangle & \langle 1|\,\hat{\mathbf{O}}\,|2\rangle \\ \langle 2|\,\hat{\mathbf{O}}\,|1\rangle & \langle 2|\,\hat{\mathbf{O}}\,|2\rangle \end{pmatrix} = \begin{pmatrix} o_1 & 0 \\ 0 & o_2 \end{pmatrix}$$

Matrix representation with operator’s eigenvectors is **diagonal** matrix

(nonzero elements (eigenvalues) on diagonal).

For general basis vectors(not eigenvectors),

matrix representation is generally **not** diagonal.

An **alternate(and useful) method** for defining operators uses the "ket" and "bra" vectors.

Consider the quantity

using our rules we get

$$\mathcal{P}_{fg} = |f\rangle\langle g| \longrightarrow \mathcal{P}_{fg}|V\rangle = (|f\rangle\langle g|)|V\rangle = \langle g\,|\,V\rangle\,|f\rangle = number \times |f\rangle$$

Thus, the new object is an **operator**! ⟶ Called the **"ket-bra"**

If g = f, then we get $\mathcal{P}_f = |f\rangle\langle f|$ = **projection** operator i.e.,

$$\mathcal{P}_f|V\rangle = (|f\rangle\langle f|)|V\rangle = \langle f\,|\,V\rangle\,|f\rangle$$

"projects" $|V\rangle$ onto $|f\rangle$ i.e., final vector now in that direction

so $\langle f\,|\,V\rangle$ **gives amount of** $|V\rangle$ **in direction of** $|f\rangle$
the component

Notice we used the rule: $(\langle f|)(|f\rangle) = \langle f\,|\,f\rangle$ **to simplify notion** $|| \Rightarrow |$
return to old braket symbol

Now consider set of orthonormal basis vectors: $\{|n\rangle\},\ n = 1, 2, 3, ..., N$

orthonormal **—>** $\langle k\,|\,n\rangle = \delta_{kn}$

We have(doing some algebra for practice):

$$(|k\rangle\langle k|)(|k\rangle\langle k|) = |k\rangle(\langle k\,|\,k\rangle)\langle k| = |k\rangle(1)\langle k| = |k\rangle\langle k|$$

Another example:

$$(|k\rangle\langle k|)(a\,|1\rangle + b\,|2\rangle) = |k\rangle(a\,\langle k\,|\,1\rangle + b\,\langle k|2\rangle) = (number\rangle \times |k\rangle$$

Now for some useful properties for later derivations and examples of use of the language.

RESULTS are important - ALGEBRA is for those learning how to do it themselves!

Consider operator $\sum_{k=1}^{n} |k\rangle\langle k|$ **—> sum of allprojection operators**

Remember from earlier that for any Ket vector, we can write (example in 2 dimensions)

$$|V\rangle = \langle 1 \,|\, V\rangle \,|1\rangle + \langle 2 \,|\, V\rangle \,|2\rangle$$

or in general

$$|V\rangle = \sum_{k=1}^{2} \langle k \,|\, V\rangle \,|k\rangle$$

so in n dimensions we have(2 terms in sum —-> n terms)

$$|V\rangle = \sum_{k=1}^{n} \langle k \,|\, V\rangle \,|k\rangle$$

clearly this is very powerful notation

We have

no change!!

$$\left(\sum_{k=1}^{n} |k\rangle\langle k|\right)|V\rangle = \sum_{k=1}^{n} \langle k \,|\, V\rangle \,|k\rangle = |V\rangle \longrightarrow \sum_{k=1}^{n} |k\rangle\langle k| = \hat{\mathbf{I}}$$

the identity operator

Result will be very important later for doing algebra!

Technical term: It is called a **representation of the identity**

Expectation Value — Very(almost most) important object in QM !!

$$\langle \hat{\mathbf{O}} \rangle = \langle V | \, \hat{\mathbf{O}} \, | V \rangle$$

Will turn out(see later) to be the average value of an observable from a set of measurements on an ensemble of identical systems all in the state $|V\rangle$. Let us see how.

First, we need an **alternative representation** of an operator

(also some simple algebra practice for those interested to do)

Only the result is important (will be used) for our later discussions……….

$\hat{\mathbf{B}} \, |b_k\rangle = b_k \, |b_k\rangle \quad k = 1, 2, 3, ..., n$ **eigenvector/eigenvalues equation**

⟶ $|b_k\rangle$ = eigenvectors of $\hat{\mathbf{B}}$ ⟶ b_k = eigenvalues of $\hat{\mathbf{B}}$

if operator is Hermitian ⟶ the set $|b_k\rangle$ are an orthonormal basis so that

$$\langle b_i \, | \, b_j \rangle = \delta_{ij} \qquad \text{and} \qquad \sum_{k=1}^{n} |b_k\rangle \, \langle b_k| = \hat{\mathbf{I}}$$

Then

$$\hat{\mathbf{B}} = \hat{\mathbf{B}}\hat{\mathbf{I}} = \hat{\mathbf{B}} \sum_{k=1}^{n} |b_k\rangle \langle b_k|$$

identity operator does not change anything

$$= \sum_{k=1}^{n} \hat{\mathbf{B}} |b_k\rangle \langle b_k| = \sum_{k=1}^{n} b_k |b_k\rangle \langle b_k|$$

operator moves through summation and operates on Ket vector

or

$$\hat{\mathbf{B}} = \sum_{k=1}^{n} b_k |b_k\rangle \langle b_k|$$

—> operator written in terms of eigenvalues/eigenvectors or projection operators

Called the "spectral representation" of the operator

Now further algebra gives

(if interested see full mathematics supplementary lecture notes for details)

$$\hat{\mathbf{B}}^2 = \sum_{k=1}^{n} b_k^2 |b_k\rangle \langle b_k| \qquad \hat{\mathbf{B}}^3 = \sum_{k=1}^{n} b_k^3 |b_k\rangle \langle b_k| \qquad \hat{\mathbf{B}}^n = \sum_{k=1}^{n} b_k^n |b_k\rangle \langle b_k|$$

Definition: Power series representation of a function:

Sometimes a function can be written as a "power series" as shown below:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots = \sum_{k=0}^{\infty} a_k x^k$$

Examples(these are the real definitions of the functions):

$$e^{\alpha x} = \sum_{k=0}^{\infty} \frac{1}{k!} \alpha^k x^k$$

$$\sin \alpha x = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} (-1)^k \alpha^{2k+1} x^{2k+1} \qquad \cos \alpha x = \sum_{k=0}^{\infty} \frac{1}{(2k)!} (-1)^k \alpha^{2k} x^{2k}$$

Now consider

$$e^{\alpha x} = \sum_{k=0}^{\infty} \frac{1}{k!} \alpha^k x^k \xrightarrow{x \to \hat{\mathbf{O}}} e^{\alpha \hat{\mathbf{O}}} = \sum_{k=0}^{\infty} \frac{1}{k!} \alpha^k \hat{\mathbf{O}}^k$$

get a function of an operator now

In special case where $\hat{\mathbf{O}}^2 = \hat{\mathbf{I}}$ (algebra is in mathematics supplement)

$$e^{i\alpha \hat{\mathbf{O}}} = \cos \alpha \hat{\mathbf{I}} + i \hat{\mathbf{O}} \sin \alpha$$

This is one of the most powerful relations in QM

(original derivation was for ordinary functions by Euler in 16th century

$$e^{i\theta} = \cos\theta + i \sin\theta \qquad \rightarrow e^{i\pi/2} = i \ , \ e^{i\pi} = -1$$

—> real definition of i

In general, if we need an operator for an arbitrary function f(x) we do the following:

if $\hat{\mathbf{B}}\,|b_k\rangle = b_k\,|b_k\rangle \quad k = 1, 2, 3, ..., n$ **then** $f(\hat{\mathbf{B}}) = \sum_{k=1}^{n} f(b_k)\,|b_k\rangle\,\langle b_k|$

Another Important Definition

$$\text{commutator } = [\hat{\mathbf{A}}, \hat{\mathbf{B}}] = \hat{\mathbf{A}}\hat{\mathbf{B}} - \hat{\mathbf{B}}\hat{\mathbf{A}}$$

ordinary numbers “commute” but operators may not “commute”

This is mathematical property behind Heisenberg uncertainty principle as we will see later.

Reminder: Useful properties to remember:

$$|V\rangle = v_1\,|1\rangle + v_2\,|2\rangle \quad , \quad |U\rangle = u_1\,|1\rangle + u_2\,|2\rangle$$

$$\langle V\,|\,U\rangle = v_1^* u_1 + v_2^* u_2 \qquad \langle U\,|\,V\rangle = u_1^* v_1 + u_2^* v_2 = \langle V\,|\,U\rangle^*$$

$$|V\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad \langle V| = \begin{pmatrix} v_1^* & v_2^* \end{pmatrix} = |V\rangle^\dagger$$

Suppose in the middle of a dream, I have an interesting thought….(this is example of how a theoretical physicist works…….)

$\langle \hat{\mathbf{B}} \rangle = \langle V | \hat{\mathbf{B}} | V \rangle$ = expectation value

$\hat{\mathbf{B}}$ is Hermitian —>

$\hat{\mathrm{B}}$ eigenvectors are complete orthonormal set = basis —> can make expansions —> so write

$$|V\rangle = \sum_{k=1}^{n} d_k |b_k\rangle \quad , \quad d_k = \langle b_k | V \rangle$$

definition of components

and

$$\langle V| = \sum_{k=1}^{n} d_k^* \langle b_k|$$

We can always do this with any basis set!!

More algebra gives(see supplement):

$$\langle V | \hat{\mathbf{B}} | V \rangle = \langle \hat{\mathbf{B}} \rangle = \sum_{k=1}^{n} b_k |d_k|^2$$

expectation value of the operator in the state |V> using eigenvector of operator **B**

But, in general, if expectation value = average value, then we must have

$$\langle \hat{\mathbf{B}} \rangle = \sum_{allowed\ values} (\text{allowed value of B}) \times (\text{Probability of that value})$$

which is just the **definition** of the average value!!

Now we just derived

$$\langle \hat{\mathbf{B}} \rangle = \sum_{k=1}^{n} b_k |d_k|^2$$

Comparing(see arrows) last two equations - we might guess that in our future theory

allowed measurement values = eigenvalues

and

probability of a value = square of corresponding vector component value

Now we can write

$$|d_k|^2 = d_k^* d_k = (\langle b_k \mid V \rangle)^* \langle b_k \mid V \rangle = |\langle b_k \mid V \rangle|^2 = \langle V \mid b_k \rangle \langle b_k \mid V \rangle$$

This assumption would then say that

$|\langle b_k \mid V \rangle|^2$ = probability of observing value b_k when in state $|V\rangle$

It also says

$$\langle V \mid b_k \rangle \langle b_k \mid V \rangle = |\langle b_k \mid V \rangle|^2 = |\text{component}|^2$$

= probability of observing value b_k in state $|V\rangle$

= expectation value of the projection operator $|b_k\rangle \langle b_k|$

These ideas will become postulates - that is what random thinking sometimes produces.

That is the kind of dream a theoretical physicist might have!

That is all mathematics we will need for now(we will add a couple of things later) to proceed with developing QM.

Lots of repetition coming…….as we use this new language!

As we worked through the mathematics

we have laid the ground work for the postulates of QM

(in some cases we have actually stated the postulate already).

We will not need any more mathematics than I have shown you already, although we will add an extension here and there.

You will get better at using the mathematics as we will use it in class and you get used to the language.

If you cannot follow all the details of the mathematics, then just follow along with the ideas that I will present based on the mathematics.

That will be sufficient to understand QM.