Lecture 1 : Wave Mechanics

Let's start with a review

Schrödinger Equation

It has the form

$$i\hbar \frac{\partial \Psi(x,t)}{\partial t} = \left(\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x,t) \right) \Psi(x,t) \qquad (1)$$

This is not the most general equation: this is just valid for the non-relativistic case, since the energy operator,

$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$, is nonrelativistic

This is 1D , not 3D . But, ok!

It is important to know that $\Psi(x,t)$, referred to as "wave function" belongs to complex numbers!

This BY NECESSITY: potential is real, but the LHS of (1) is complex! It MUST be complex!

The fact that Ψ is complex is really fundamental!

In other subjects, this property was present, but not so strong.

In electromagnetism, for example, the electric field $\underline{E}(x,t)$ for a circularly polarized way is:

$$\underline{E}(z,t) = \mathrm{Re}\left\{ E_0 (\hat{x} + i\hat{y}) e^{i(kz-\omega t)} \right\} \quad \hat{x},\ \hat{y} \text{ unit vectors}$$

an electric field is REAL, but its computation involves complex numbers.

As well, Maxwell equations are REAL!

Complex numbers are just a neat way to write a complicated, but real, EM field. All "i"'s are auxiliary!

Remarks about (1) (Sch. Eqn)

1) It is a first order differential equation in time. This means that if we know the wavefunction in all space for a certain t, we can compute its evolution in time! Its values at successive times!

So,

$\Psi(x, t_0)$ at one t_0 determines Ψ at <u>all</u> times

2) It is linear: if given 2 solutions Ψ_1, Ψ_2, then $a_1\Psi_1 + a_2\Psi_2$, $a_1, a_2 \in \mathbb{C}$, is a solution. Superposition principle!

For a generic complex number z,
$z = a + ib$, $z^* = a - ib$ (conjugate), $|z| = \sqrt{a^2+b^2} = \sqrt{zz^*}$. $a, b \in \mathbb{R}$

What makes this equation be physical and not just math is the interpretation of the wavefunction as a probability density! Let

$P(x,t)$ be (def) $g(x,t)$,

$g(x,t) P(x,t) = \Psi^*(x,t)\,\Psi(x,t)$ [$\Psi^*(x,t)$ is complex conjugate], [probability density]

This means that if we take $P(x,t)$ and multiply times a dx,

$P(x,t)\,dx$,

this is the probability to find a particle in the interval $[x, x+dx]$, at time "t".

This is a postulate that gives physical meaning to the wavefunction! So, since the particle must be somewhere, (and we are describing just 1 particle),

$\int_{-\infty}^{+\infty} |\Psi(x,t)|^2\,dx = 1$, $\forall t$ particle is somewhere in space!

Notice that $[\Psi] = \frac{1}{\sqrt{L}}$, L meter!

Suppose we have a wavefunction such that $\int_{-\infty}^{+\infty} |\Psi(x,t_0)|^2\,dx = 1$.
Is it true that, for <u>all</u> times, given this hypothesis (integral of $P(x,t)$ equal 1 for a $t=t_0$), the integral remains equal to 1? If not, we could have problems! In fact we said that knowing the wavefunction at some time corresponds to knowing it at any time! So, if it is a "good" (normalizable) at $t=t_0$, is it also at other times?
Yes, it is!
Exercise: show that

$\frac{d}{dt}\int_{-\infty}^{+\infty} |\Psi(x,t)|^2\,dx = 0$

The integral might depend on time. Is it true? If $\frac{d}{dt} = 0$, then it does not change in time! [Check 8.04 notes!]
The procedure is: differentiate Ψ, Ψ^*, integrate by parts, and throw away the "boundary terms" at $\pm\infty$. This can be done because if the wavefunction does not go to 0 at ∞, it is NOT normalizable! =)
But also the derivative should not blow up! But, go to 0!
We will treat <u>mild</u> potential functions!

Notice that we are possibly considering time-dependent potentials! Sometimes it might be necessary!

Probability current
This is a $J(x,t)$ given by

$J(x,t) = \frac{\hbar}{m}\,\mathrm{Im}\left\{\Psi^*\frac{\partial\Psi}{\partial x}\right\}$,

and this goes together with this probability density, $P(x,t)$.
In electromagnetism we have currents and charge densities: a vector $\underline{J}$, and ρ. The most famous thing related to them is the conservation law,

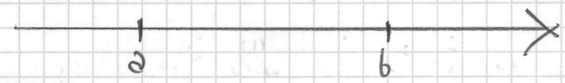
$\vec{\nabla}\cdot\underline{J}(x,t) + \frac{\partial\rho}{\partial t} = 0$

or, in 1D,

$\frac{\partial J}{\partial x} + \frac{\partial\rho}{\partial t} = 0$

this <u>means</u> charge conservation! Charge is never destroyed or created! If we consider a volume, charge could enter or escape, but not be created or destroyed!

We can interpret in 1D as: the probability to find a particle in $[a,b]$ is



$P_{ab}(t) = \int_a^b |\Psi(x,t)|^2\,dx$

it is possible to find that [review]

$\frac{dP_{ab}}{dt} = J(a,t) - J(b,t)$

So the probability changes depending on how current leaks from a and b!

Last thing for wavefunction.
Two wavefunctions Ψ_1 and Ψ_2 are said to be "equivalent" if
$\Psi_1(x,t) = \alpha\, \Psi_2(x,t)$, $\alpha \in \mathbb{C}$.

Here, if $|\alpha| = 1$, and then it is just a phase, well, Ψ_1 and Ψ_2 have the same physical meaning!
This puts some constraint about how we define observables! An observable should satisfy this property! If we have Ψ_1 or Ψ_2, the observable should not change!

Now, we reviewed Schrödinger equation. Let's then review its most important solutions: energy eigenstates, stationary states, and so on.

Stationary solutions

Up to this moment, we used capital psi, Ψ. We want to distinguish it from another!
Assume that the potential $V(x,t) = V(x)$: time independent!
In this case,

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi ,$$

where

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \qquad \text{(at this point)}$$

This $\hat{H}$ is an operator acting to the right. What we mean is that this operator acts on some space, so takes elements of that space and "moves them around" in the space.
$\hat{H}$ is called "Hamiltonian operator" and it is time-independent.
What is a stationary state?
A stationary state of energy $E \in \mathbb{R}$ is a wavefunction $\Psi(x,t)$ of following form:

$$\Psi(x,t) = e^{-i\frac{Et}{\hbar}} \psi(x)$$

ψ is NOT capital! it's time independent!

$\psi \in \mathbb{C}$, but does not depend on time!

If we compute expectation values of observables for stationary states, they will be time-independent!!!

For instance, probability density! In fact,

$$J = \frac{\hbar}{m} \operatorname{Im}\left\{ \Psi(x,t) \frac{\partial}{\partial x} \Psi^*(x,t) \right\},$$

but then by substituting $\Psi \to e^{-i\frac{Et}{\hbar}}\psi(x)$, the exponentials disappear and all the remaining part is time-independent!

Plugging Ψ in (1),

$$\Rightarrow i\hbar \left(\frac{\partial}{\partial t} e^{-i\frac{E}{\hbar}t} \right)\psi(x) = e^{-i\frac{E}{\hbar}t} \hat{H}\psi(x)$$

then, by simplifying,

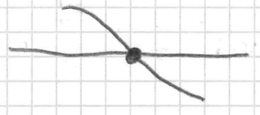
$$\to E\psi(x) = \hat{H}\psi(x)$$

this is the time-independent Schrödinger equation!

We could say a lot of things about this equation!
Time has been factored out. Moreover, this seems real! So also $\psi(x)$ can be real!
This is a 2nd-order differential equation in space!

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x) = E\psi(x) \qquad (2)$$

So if we know the wavefunction and its derivative at one point (boundary condition), we can know ψ everywhere!
Like, $\psi(x_0)$, $\psi'(x_0)$ suffices for a solution when $V(x)$ is regular (not too strange/complicated).
So, if $\psi(x_0)=0$, $\psi'(x_0)=0$, then $\psi(x)=0$ (for regularity)!
It's a property of this ODE!
The wavefunction should go to 0, slowly!



What is an energy eigenstate?
It is a solution of (2)! ψ!
And, the set of values of E is the spectrum!
If we have more than one ψ associated to the same E, then

the spectrum is said to be "degenerate".
Degeneracies are consequence of strange, interesting physics!

We allow potentials that:
- feature discontinuity (square well!)
- failure to be bounded! (harmonic oscillator!)
- it can include δ functions!
(but not derivatives nor powers of δ! =c)

Potentials are no more complicated than that!

$\psi(x)$ is continuous and bounded, and its derivative is bounded!

Let us not focus on normalizability. For instance, momentum eigenstates are useful, but NOT normalizable

Properties of the eigenstates
There are two types of identities that we should be aware of.
The eigenstates of $\hat{H}$ can be "listed" by their energies: $E_0, E_1, E_2 \dots$
... and corresponding $\psi_0, \psi_1 \dots$, so that

$\hat{H}\psi_n = E\psi_n$

this, when we have reasonable potentials
We can have orthonormal ψ_n! This, in the sense that we will consider functions as sort of "vectors of an infinite dimensional vector space". Not just in 2 or 3 dimensions or so. Why? Well, think of a function as an ∞ array of numbers! Just like we define a vector with its components! Like, then, the "first vector" is "directed" along one direction, and so on.
Saying that two functions are orthonormal can be done by defining an extension of the concept of "dot product":

$$\int_{-\infty}^{+\infty} \psi_m^*(x)\,\psi_n(x)\,dx \overset{?}{=} \delta_{mn}$$ orthogonality

If this integral equals δ_{mn}, the two functions are said to be "orthonormal"!

Another subtle property of the $\{\psi_n\}$ is that the set of $\{\psi_n\}$ is enough to expand, to represent, any function in this space! The space of wavefunctions! We can write any wavefunction as their superposition! And this, is provided just by the solution of the stationary Sch. eqn

Orthonormality and Completeness!

$$\psi(x) = \sum_{n=0}^{\infty} b_n \psi_n(x), \qquad b_n \in \mathbb{C}.$$ Completeness

Why is this so useful? Well, let's assume that we have, at a certain time, a wavefunction, $\Psi(x, t=0)$.

Thanks to completeness, we can write it as

$$\Psi(x, t=0) = \psi(x) = \sum_n b_n \psi_n(x) \qquad (3)$$

If we can calculate the coefficients $\{b_n\}$ of the linear combination, we can solve the problem of time evolution, because

$$\Psi(x,t) = \sum_n b_n e^{-i\frac{E_n}{\hbar}t} \psi_n(x)$$!! We evolve each energy eigenstate

Of course, we have to compute the b_n.
This can be done by using orthonormality:

$$\psi(x) = \sum_n b_n \psi_n(x)$$

$$\int_{-\infty}^{+\infty} \psi_m^*(x)\,\psi(x)\,dx = \sum_n b_n \int_{-\infty}^{+\infty} \psi_m^* \psi_n\,dx = \sum_n b_n \delta_{mn} = b_m \quad !$$

$$b_m = \int_{-\infty}^{+\infty} \psi(x)\,\psi_m^*(x)\,dx$$

Let us plug back b_m in (3):

$$\psi(x) = \sum_{n=0}^{\infty} \left(\int_{-\infty}^{+\infty} \psi_n^*(x')\psi(x')\,dx' \right) \psi_n(x)$$

We gain something if we write ψ as equal to something, times ψ!

Let's assume that everything converges. So,

$$= \int_{-\infty}^{+\infty} dx' \sum_{n=1}^{\infty} \left(\psi_n^*(x') \psi_n(x) \right) \psi(x') = \dots$$

This way of writing is nice, because we can learn something from it.

This must be a very peculiar function if, integrated times ψ, returns ψ again! What could it be?

Well, something like

$$\dots = \int dx' \, K(x,x') \psi(x')$$

K is a function is a set of function that takes the value at x!

A δ function!

The claim is:

$$\sum_{n=1}^{\infty} \psi_n^*(x') \psi_n(x) = \delta(x-x')$$

This must be true, if we can expand any function in terms of the eigenfunctions! This is the completeness!

We have two sums in there expressions: a sum over space, x, which is the integral, and the sum over indices, over labels.

The othogonality relation

$$\int_{-\infty}^{+\infty} \psi_m^*(x) \psi_n(x) \, dx = \delta_{mn}$$

keeps fixed the labels and sums over space.

The completeness relation

$$\sum_n \psi_m^*(x') \psi_n(x) = \delta(x-x')$$

sums over labels and space is "fixed"! It is some sort of "identity matrix" in x, x'! x, x' are like indices! The same idea as δ_{mn}

There is some duality between E and x!

Orthonormality and completeness are somehow dual!

Expectation values

Given a time-independent operator, $\hat{A}$, then we have the expectation value of it on a state, Ψ

$\langle \hat{A} \rangle_\Psi$ (t) (It is in general expected to be a function of time!)

This is defined as:

$$\langle \hat{A} \rangle_\Psi = \int_{-\infty}^{+\infty} dx \, \Psi^*(x,t) \left(\hat{A} \Psi(x,t) \right),$$

where Ψ is assumed to be a normalized state!

If Ψ is stationary time exponentials disapear due to the complex conjugation. $\hat{A}$ is time-independent. But ...

... if this state is nonstationary, we might have time dependence.

Let's study the expectation value of $\hat{H}$!

$$\langle \hat{H} \rangle_\Psi = \int_{-\infty}^{+\infty} dx \, \Psi^* \left(\hat{H} \Psi(x,t) \right)$$

$\Psi = \sum_n b_n \psi_n$ (substitute this)

(look notes) ...

$$= \sum_n |b_n|^2 E_n$$

it is this! This is time-independent!

This is a special property of the Hamiltonian operator: energy is conserved!

Normalized

If we have a wavefunction Ψ that is not normalized, we may want to do it:

$$\Psi \longrightarrow \frac{\Psi(x,t)}{\sqrt{\int \Psi^* \Psi \, dx}}$$ this! is!

So, for $\hat{A}$, the exp. value is

$$\langle \hat{A} \rangle_\Psi = \frac{\int_{-\infty}^{+\infty} dx\, \Psi^*(x,t)\, \hat{A}\, \Psi(x,t)}{\int_{-\infty}^{+\infty} \Psi^*(x,t)\, \Psi(x,t)\, dx}$$

thus if Ψ is not normalized!

Lecture 2: Wave mechanics (continued)

Last time we discussed about energy eigenstates. A bound state is something, a state, which is not "spread" all over the space.
"Bound states" are always energy eigenstate: it is a term aimed at identifying some peculiar energy eigenstate. An energy eigenstate may or may not be a bound state.
Energy eigenstates: time independent wavefunctions.
In 1 dimension, an energy eigenstate $\psi(x)$ is a bound state if

$$\psi(x) \to 0 \quad \text{when } |x| \to \infty$$

This is just part of the full wavefunction. We still lack a phase, containing information about the energy.
Energy eigenstate eq. is:

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}\left[E - V(x)\right]\psi(x) = 0 \qquad (1)$$

In order to clean up this equation we can manipulate the constants.
Let's scale the energy, defining a adimensional $\mathcal{E}$

$$\mathcal{E} = \frac{2m}{\hbar^2}E, \quad U(x) = \frac{2m}{\hbar^2}V(x),$$

$$\hookrightarrow \psi'' + (\mathcal{E} - U(x))\psi = 0 \qquad (2)$$

$\mathcal{E}$, U are normalized energy and potential

Let's discuss some theorems some results. Proofs can be found in the notes!

Theorem 1
If we talk about bound states in 1 dimension, there are no degeneracies.
"There is no degeneracy for bound states of 1 dimensional potentials."
How can we prove it? By contradiction: if we assume that there is a degeneracy, so $\psi_1 \neq \psi_2$ different, with $\mathcal{E}_1 = \mathcal{E}_2$, by solving (2) and doing some algebra, we can prove that one solution must equal the other, times a constant!

Theorem 2
"Energy eigenstates $\psi(x)$ can be chosen to be real"
Proof strategy: from a complex solution $\psi(x)$, this implies the existence of 2 real solutions, DEGENERATE. This can be shown by finding a solution, and proving that its conjugate, ψ^*, still is a solution. Then,

$$\psi_r(x) = \frac{1}{\sqrt{2}}\left[\psi(x) + \psi^*(x)\right) \qquad \psi_i = \frac{1}{i\sqrt{2}}\left(\psi - \psi^*\right)$$

So: a complex solution implies two real solutions, we can choose solutions to be real.

Corollary, for 1D bound states
If we're talking bound states of 1D potential, EVERY solution is real.
In fact, ψ_r and ψ_i are equal, due to Theorem 1

"For bound states of one-dimensional potentials, any solution is, up to a phase, equal to a real solution".
"By Th 1, since the two ψ_r, ψ_i are degenerate,"

$$\psi_r = c\,\psi_i$$

since ψ_r, ψ_i are real, then the constant "c" must be real. Therefore, even if we consider

$$\psi = \psi_r + i\psi_i,$$

it is equal to

$$= \underbrace{(1+ic)}\,\psi_r(x)$$

↳ and this is a phase!

Theorem 3

If the potential $V(x)$ is even, that is $V(-x)=V(x)$, the eigenstates can be chosen to be even or odd under the transformation $x \to -x$

So. The word "chosen" is very important.

Strategy for proof. Begin with a $\psi(x)$ (which is neither even nor odd. You can see that both $\psi(x)$ and $\psi(-x)$ are solutions, with the same energy. Once that we prove this point, and since Th. 2 guarantees that the wavefunctions can be real, we can choose

$\psi_{sym} = \frac{1}{\sqrt{2}}\left(\psi(x)+\psi(-x)\right)$, $\to$ even

$\psi_{asym} = \frac{1}{\sqrt{2}}\left[\psi(x)-\psi(-x)\right]$ $\to$ odd

Corollary

For bound states in 1D, we can delete the word "chosen": ALL the solutions are either odd or even! $\Rightarrow$ Not just "choose"!

Proof strategy: by using previous results, and the ψ_{sym}, ψ_{asym}, degenerate, using Th. 1,

$\underbrace{\psi(-x)=c\,\psi(x)}_{sym}$ and $\underbrace{\psi(x)=-c\,\psi(x)}_{asym}$!

The bound states of a 1D even potential can be either even or odd, NOT arbitrary.

Only bound states can be normalized!

Nature of the spectrum

We are going to discuss the nature of the spectrum.

Let's go back to (1) and rewrite it like

$$\psi'' = -\frac{2m}{\hbar^2}\left(E-V(x)\right)\psi \qquad (3)$$

We always "assume" that ψ is continuous. Indeed, we stated that we don't want singularities worse than δ functions in potential.

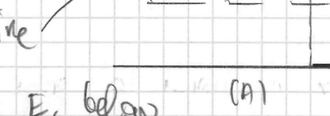
So the $V(x)$ can contain at most a δ function but, for (3), it means that at most the 2nd derivative of ψ can contain a δ. But then, if ψ is not continuous, it is impossible. ψ must be continuous.

If ψ is continuous, several possibilities for V are considered:

POSSIBILITIES

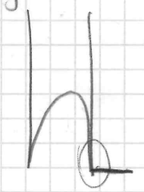
1. V is continuous. So, ψ'' is continuous, as $(E-V(x))\psi(x)$ is continuous. Then also ψ' is continuous.

2. V has finite jump discontinuities. In this case, the product $V(x)\psi(x)$ has finite jumps. But then, ψ'' has finite jumps. So, ψ' is continuous, with slope discontinuity!



3. V has δ functions.
In this case, ψ'' has a δ function, then ψ' "jumps"!
So, the wavefunction $\psi(x)$ has a corner!

4. V has a "hard wall": potential suddenly goes to ∞!
This occurs in the infinite well, where the wavefunction goes to 0, but a discontinuity of first derivative occurs.

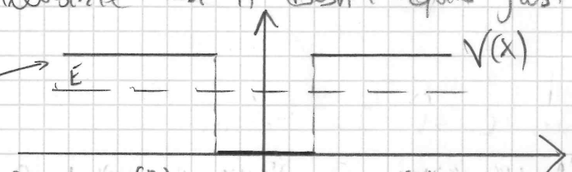


To summarize this,

{ψ and ψ' are continuous unless $V(x)$ has δ functions or hard walls, in which case ψ' can have finite jumps.}

Let us try to get more intuition about the Theorem 1, not because the proof is not solid. But we can illustrate that it doesn't require just math.

Let us assume a potential like



and look for an energy like E_1 below the barrier.

What we are going to see is an elaboration of what we can read in "Shankar".

Why we would get, for this situation, no degeneracies? Why is the spectrum quantized?

Let us try to write the solutions and study their parameters, and how many conditions they must satisfy.

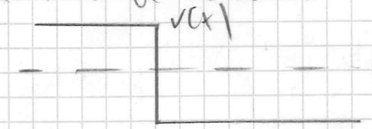
In region (A), the wavefunction is a decaying exponential, something like αe^{Kx}, $K>0$. We need "α" to determine the solution

In (B) the solution is oscillatory, so $a\cos(x)+b\sin(x)$: 2 coefficients a and b. In (C), the solution is of an exponential type,

$$\beta e^{-kx}$$

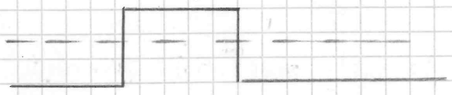
We have 4 parameters: α, β, a, b
Given that we can multiply the solution by a constant and it still solves Schrödinger, we have 3 constants to fix!
But, ψ and ψ' must be continuous, as the potential is nice enough, see previous discussion. So, 4 conditions!
3 parameters (4 - 1 per normalization), 4 conditions!
BAD!
But we are missing something: K! K is related to energy! The fourth parameter is the energy!
The solution exists, for some energy! Not for all values!
Only some energies are ok!
This leads to the fact that the spectrum must be discrete!

Other situation: the potential step



For this we get a continue spectrum,
nondegenerate. For every energy we have a unique solution.

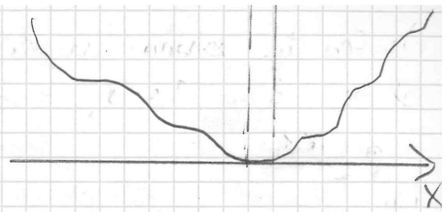
Last case: barrier



We have a continue spectrum,
doubly degenerate.

All of this is shown in the notes!

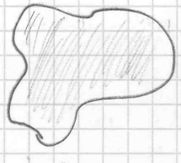
Another result, a theorem, hard to prove (not proved in notes, too hard):
the nodes theorem
Theorem
Given the discrete bound state spectrum of a 1D potential, if we list the energies, $E_1 < E_2 < E_3$, where E_1 is the ground state energy, then the eigenstates $\psi_1(x)$, $\psi_2(x)$ are such that ψ_1 has no nodes, no intersections with 0, ψ_2 has 1 node, ψ_3 has 2 nodes, and ... : ψ_n has $n-1$ nodes!
There are several arguments to try to show this. Physicists have a rather delicate argument, which is based on approximating any potential by infinite square wells, to begin. For instance,



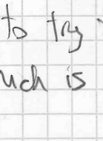
the dashed line is the ∞ square well approximating the potential in a neighborhood.
We can make this "window" wider.
Since we know that the first wavefunction of the ∞ square well has 0 nodes, we change this "screen", this window, and the wavefunction cannot acquire nodes.
Moreover, it can be shown (problem set) that the successive solution (in terms of energy) always has one more node.

Now, let's study something new, for the rest of the lecture

The variational problem
This problem has to do with calculus of variations, which is a rather complex subject.
This calculus of variation is a more complex version of the problem of Calculus I: to find maxima and minima of a function. In that case one takes the derivative of a function, sets it equal to zero and finds the position of the max/min.
The variational problem is a situation where we want to maximize/minimize something, but this is performed not fixing a function and working on it, but finding a function that gives the min/max of this "something". The fact is: the unknown of the variational problem is not a point: it is a function!
These kinds of problems where first discussed by Newton:



given a cross-section like this, imagine to make a solid out of it, by tapering it and ending with some "hat", some "vertex", and make it flow through a viscous fluid.
How can we make the resistance from the viscous fluid minimum?
We need the SHAPE of tapering that minimizes resistance.
Apparently Newton solved the problem but didn't explained how.
Then, Bernoulli in 1696 (Johann Bernoulli), poses the problem of:
given two points in the plane, asked to find the curve of shortest time for fall, from A to B.



6th June 1696 Leibniz received this problem, 16 June answered with a complete solution.

Apparently Johann Bernoulli proposed this problem just to show that his older brother, Jacob, was not able to solve it =D
Johann is the father of Daniel, the guy of fluid dynamics stuff.
This problem is not so easy, and it can be solved by calculus of variations.
As the Bernoulli problem was aimed to find a function that minimizes time, Schrödinger equation, or better its eigenfunctions, minimize something else!
We want to solve the problem

$$\hat{H}\psi = E\psi(\vec{x})$$

this is valid for more than 1 dimension.
Firstly, before understanding what calculus of variation is useful for, in this scenario. Let's learn something more about the ground state energy of this system.
Ground state energy
Let us consider an arbitrary $\psi(\vec{x})$, normalized. So,

$$\int d\vec{x}\, |\psi(\vec{x})|^2 = 1$$

$\psi(\vec{x})$ is not necessarily a solution of Schrödinger equation! it is an arbitrary, normalized function.
Let us compute nevertheless the following quantity:

$$\int d\vec{x}\, \psi^*(\vec{x})\, \hat{H}\, \psi(\vec{x}) \qquad \left(\equiv \langle \hat{H} \rangle_\psi\right)$$

The remarkable fact about this quantity, is that it provides an upper bound for the ground state energy!

$$E_{gs} \leq \langle \hat{H} \rangle_\psi$$

and ψ is ARBITRARY!
It would be ideal if we also had a lower bound!
So, if we try all functions that may look like the ground state wavefunction, we get numbers, that bound E_{gs}.
$\psi(\vec{x})$ is called "trial wavefunction".

"Proof" of this theorem

Assumption: let's assume that we don't have a continuous spectrum.

This is useful to write a simpler proof.

$$E_{gs} = E_1 \leq E_2 \leq E_3 \cdots$$

we may even consider degeneracies (more or equal than 2D).
and,

$$\hat{H}\psi_n = E_n \psi_n$$

ψ_n is a trial wavefunction.

By completeness, our $\psi(\vec{x})$ can be expanded as a series/superpositon of energy eigenstates:

$$\psi(\vec{x}) = \sum_{n=1}^{\infty} b_n \psi_n(\vec{x})$$

$\psi(\vec{x})$: trial: DOES NOT SOLVE SCH. EQN
$\psi_n(\vec{x})$: energy eigenstates!

this is just a representation!

$$\int |\psi(\vec{x})|^2 dx = 1$$

$$= \sum_{n=1}^{\infty} |b_n|^2 \quad ! \text{ (by orthonormality!)}$$

Moreover,

$$\int \psi^* \hat{H} \psi \, d\vec{x} = \sum_{n=1}^{\infty} |b_n|^2 E_n$$

Notice that $E_2, E_3, E_4, \ldots$ are all bigger than E_1 (or at most equal)
So, we can say that:

$$\sum_{n=1}^{\infty} |b_n|^2 E_n \geq \sum_{n=1}^{\infty} |b_n|^2 E_1 = E_1 \sum_{n=1}^{\infty} |b_n|^2 = E_1 \;=)$$

And this proves our theorem =)
So, how is this related to the variational principle?
Firstly: if $\psi(x)$ is not normalized, we know that

$$\frac{\psi(x)}{\sqrt{\int |\psi(x)|^2 dx}}$$ is normalized =)

So,

$$E_{gs} \leq \frac{\int \psi^* \hat{H} \psi \, d\vec{x}}{\int \psi^* \psi \, d\vec{x}}$$

and we don't need to work with normalized wavefunctions =)

Let $\mathcal{F}[\psi]$ be

$$\mathcal{F}[\psi] = \frac{\int \psi^* \hat{H} \psi \, d\vec{x}}{\int \psi^* \psi \, d\vec{x}}$$

$\mathcal{F}[\psi]$ is what is called a "functional": it is an expression whose input is a function and its output a number!
Just like the brachistochrone problem, the ground state wavefunction is the wavefunction that minimizes this functional!
And the functional equals E_{GS}!

A functional is a sort of function in an infinite dimensional space!

If we modify the ground state wavefunction like making it fatter or slimer somewhere, energy will increase: the bound state is the one that minimizes energy!

Actually, every single eigenstate is a critical point of this functional: a "saddle point"! (the minimum is ground state).

Example

$V(x) = -\alpha \delta(x), \quad \alpha > 0$

$E_{GS} = -\frac{m\alpha^2}{2\hbar^2}$ (known from 8.04 course)

We could try to put some Gaussian, like

~~$\psi(x) = e^{-x^2}$~~

But no, it should be better to put some more constant to make the minimization work better!

$\psi(x) = e^{-\frac{1}{2}\beta^2 x^2}$ trial wavefunction

If we compute $\mathcal{F}[\psi]$, we get a function of β. So, β can be adjusted to get the bound state!
Note that

$$\int |\psi|^2 dx = \frac{\sqrt{\pi}}{\beta}$$

So!

$$\frac{\int \psi^* \hat{H} \psi}{\int \psi^* \psi \, dx} = \frac{\beta}{\sqrt{\pi}} \int dx \, e^{-\frac{1}{2}\beta^2 x^2} \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - \alpha \delta(x) \right] e^{-\frac{1}{2}\beta^2 x^2}$$

we can solve it by MATLAB / Mathematica. However, the δ is easy, the other is more complex, but, by integrating by parts,

$$= -\frac{\beta\alpha}{\sqrt{\pi}} + \frac{\beta}{\sqrt{\pi}} \frac{\hbar^2}{2m} \int dx \left[\frac{d}{dx} e^{-\frac{1}{2}\beta^2 x^2} \right]^2 \quad \text{and this is easier.}$$

δ part [we can show first this is 0]

$$= -\frac{\beta}{\sqrt{\pi}} \alpha + \frac{\beta^2 \hbar^2}{4m}$$

This is a function of β. We want to find its minimum: the best upper bound!

$$\frac{\partial}{\partial \beta} \left[-\frac{\beta}{\sqrt{\pi}} \alpha + \frac{\beta^2 \hbar^2}{4m} \right] = -\frac{\alpha}{\sqrt{\pi}} + \frac{\beta \hbar^2}{2m} = 0$$

$$\hookrightarrow \beta = \frac{\alpha}{\sqrt{\pi}} \frac{2m}{\hbar^2}$$

$$\hookrightarrow E_{GS} \leq -\frac{m\alpha^2}{\pi \hbar^2} = \frac{2}{\pi} \left(-\frac{m\alpha^2}{2\hbar^2} \right)$$

this is the best guess we can get using a Gaussian trial function!!!

Lecture 3: Wave mechanics and Stern-Gerlach experiment

Position and momentum

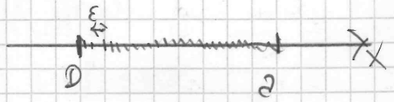
Position and momentum, in quantum mechanics, are <u>operators</u>: $\hat{x}$, $\hat{p}$.
Position and momentum operators don't commute, and

$$[\hat{x}, \hat{p}] = i\hbar \qquad (1)$$

Up to this moment we discussed a lot about wavefunctions; wavefunctions, indeed, represent the dynamics of our system. Dynamics means how our particles move in time. But time, as we are studying, is some sort of "spectator", an "arena" where things happen; however, <u>operators</u> are going on, without reference of time. We take the wavefunction, expand as a linear combination of energy eigenstates and evolve each of them.
So, what we are doing, is, basically, we are studying wavefunctions with no time dependence, $\psi(x)$.
The fact of writing the wavefunction $\psi(x)$ as a function of x, of space, means that we are <u>representing</u> it in the "x" representation: the position representation. This means that it is easy to understand how the position operator $\hat{x}$ acts on $\psi(x)$: by applying $\hat{x}$ to $\psi(x)$, "$\hat{x}\psi(x)$", we obtain another function, such that

$$\hat{x}\psi(x) \triangleq x\psi(x) \qquad \text{DEFINITION!}$$

this is a definition: $\hat{x}$ applied to $\psi(x)$ is equal to x, the position, times $\psi(x)$.
A recurrent theme in quantum mechanics is that of wavefunctions, which we will also call "states", or "vectors" as we will formally think to wavefunctions as vectors of an ∞ dimensional space.
By invoking an analogy with linear algebra, if operators act on wavefunctions, then, operators are the counterparts of <u>matrices</u>, just like wavefunctions and vectors!
So, how do we think of $\hat{x}$ or $\psi(x)$ as matrices or vectors? Well, a not-so-formal, but clear idea, is: take a wavefunction and assume that we are interested to it in the interval $[0, 2]$. We can divide this interval in several segments, and give the values for each point, as if they were the components of a vector!



Given N intervals, each having width equal to ε, we can think of $\psi(x)$ as a vector as a vector with N components

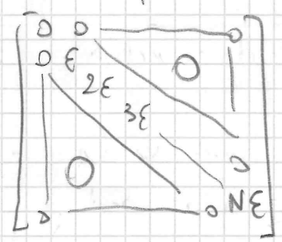
$$\psi(x) \text{ on } [0, L] \longleftrightarrow \begin{bmatrix} \psi(0) \\ \psi(\varepsilon) \\ \psi(2\varepsilon) \\ \psi(3\varepsilon) \\ \vdots \\ \psi(N\varepsilon) \end{bmatrix} \qquad N\varepsilon = a$$

$N+1$ column vector!

Depending on how much accuracy we want, we can choose ε small enough.
In this strange, approximated way of thinking, which is the role of the operator/matrix $\hat{x}$?

$\hat{x} \longleftrightarrow$ $(N+1)\times(N+1)$ matrix, and its entries are 0 everywhere, except on the diagonal, where they are.

$\hat{x} \longleftrightarrow$



this, we claim, is the way how we should think of the $\hat{x}$ operator
Let us check it!!!

$$\hat{x}\psi(x) \longleftrightarrow \begin{bmatrix} 0\psi(0) \\ \varepsilon\psi(\varepsilon) \\ 2\varepsilon\psi(2\varepsilon) \\ \vdots \\ N\varepsilon\psi(N\varepsilon) \end{bmatrix}$$

diagonal matrix times a vector: easy to show!
!!!
And this looks like $x\psi(x)$!

So, once that we know how the operator $\hat{x}$ is defined and we understood it, we could also find its expectation value!

$$\langle \hat{x} \rangle_\psi = \int dx\, \psi^*(x)\left(\hat{x}\psi(x)\right) = \int dx\, \psi^*(x)\, x\, \psi(x)$$

Now, question: are there eigenstates of the position operator?
Yes! But they are a bit "singular" $\Rightarrow$
Well, intuitively, it is something that must have a definite value of the position, something that exist only in one position: a δ function!
Let us define $\psi_{x_0}(x)$, labeled by x_0:

$\psi_{x_0}(x) = \delta(x-x_0)$

we claim that this is an eigenstate of $\hat{x}$. In fact,

$$\hat{x}\,\psi_{x_0}(x) = \hat{x}\,\psi_{x_0}(x) = x\,\psi_{x_0}(x) = x\,\delta(x-x_0) = x_0\,\delta(x-x_0) = x_0\,\psi_{x_0}(x)$$

x_0 is a single value, not a function: eigenvalue! So, if $\hat{x}$ acts on δ, it reproduces the same function, times a number: this is the definition of an eigenstate! $\delta(x-x_0)$ is an eigenstate of $\hat{x}$ with eigenvalue x_0.

Why we said "singular"? Well, it is not normalizable, so it cannot represent a particle! :-c

Position becomes even more interesting when we define its dual quantity: momentum!

Momentum is an operator as well, and it must be defined. A shorthand from 8.04 is:

$$\hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial x}$$

this means that, in the position representation, where wavefunctions are expressed as functions of x, momentum is written in this way

Why? Well, in 8.04 we said that this comes from de Broglie, the wavelenght of a wave is related to momentum, particles are waves, and so on.

So, knowing this,

$$\hat{p}\,\psi(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x)$$

Is this definition consistend and/or does it imply the commutation relation (1)? Well, we test the operators on a $\psi(x)$, a wavefunction!

$$[\hat{x},\hat{p}]\psi(x) = ?$$

This is an operator equation! Operator equations are verified by testing them on an arbitrary wavefunction! Proof:

$$[\hat{x},\hat{p}]\psi = (\hat{x}\hat{p} - \hat{p}\hat{x})\psi = \hat{x}\hat{p}\psi - \hat{p}\hat{x}\psi = \hat{x}\frac{\hbar}{i}\frac{\partial\psi}{\partial x} - \hat{p}\,x\psi =$$

$$= x\frac{\hbar}{i}\frac{\partial\psi}{\partial x} - \frac{\hbar}{i}\frac{\partial}{\partial x}(x\psi) = x\frac{\hbar}{i}\frac{\partial\psi}{\partial x} - \frac{\hbar}{i}\psi - \frac{\hbar}{i}x\frac{\partial\psi}{\partial x} = +i\hbar\psi$$

So, it is always true, without hypotheses on ψ, which can be arbitrary.

Since

$$[\hat{x},\hat{p}]\psi = i\hbar\psi \quad \forall\psi$$

then I can write $[\hat{x},\hat{p}] = i\hbar$.

Another nice thing to do could be: or, we have a nice representation of $\hat{x}$ as a matrix; do we have also one for $\hat{p}$?

We could do this by discretizing and obtaining a finite differences representation!

For example,

$$\begin{bmatrix} \psi'(0) \\ \psi'(\varepsilon) \\ \psi'(2\varepsilon) \\ \\ \\ \psi'(N\varepsilon) \end{bmatrix} \approx \frac{1}{\varepsilon} \begin{bmatrix} \psi(\varepsilon)-\psi(0) \\ \frac{\psi(2\varepsilon)-\psi(0)}{2} \\ \\ \\ \\ \end{bmatrix}$$

and such transformation can be easily written in matrix form: finite differences matrix!

And, you could also try to see if these matrices commute! For some example of wavefunction!

Or, still discussing momentum, which are its eigenstates? e^{ipx}!

With some convenient normalization,

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\,e^{ip\frac{x}{\hbar}}$$

In fact,

$$\hat{p}\,\psi_p(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}\psi_p = p\,\psi_p(x)$$

this is the eigenstate of momentum operator, and has eigenvalue "p"

What is the use of this? Well, in addition to the position representation, which considers the wavefunction as a function of x, there is the momentum representation of the wavefunction, where it is as function of p.

The relation between the two representations is given by the Fourier transform. This means that $\psi(x)$ can be represented as a "sum" of momentum eigenstates. So, for Fourier's theorem,

$$\psi(x) = \int_{-\infty}^{+\infty} dp\,\frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\,\tilde{\psi}(p) \quad (2)$$ this defines $\tilde{\psi}(p)$!

You can also invert it, as

$$\tilde{\psi}(p) = \int_{-\infty}^{+\infty} dx\,\frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}}\,\psi(x) \quad (3)$$

Let us ponder about these equations! (2) is the "sum", the "linear combination" of momentum eigenstates, where $\tilde{\psi}(p)$ is the "coefficient" for each element, quantifying the "contribution" of each momentum eigenstate
(3) is the opposite!
Let us think of this as of a change of representation: all the information about the system is contained in $\psi(x)$! But also in $\tilde{\psi}(p)$, as it contains the same information: they are just different ways of encoding the same information!
The relation between them is: we thought of $\psi(x)$ as a vector of an ∞ dimension space, in position space.
$\tilde{\psi}(p)$ could be thought in the same way, with momentum!
Between the 2, there is an integral: Fourier!

$$\underline{\tilde{\psi}}_p = \underline{\underline{M}}_{px}\,\underline{\psi}_x$$

$\underline{\underline{M}}_{px}$ is just like a change of basis! You sum over index "x" and go to index p!
Two different vectors, containing the same physics encoded in different ways!
We can do one more thing: act, with $\frac{\hbar}{i}\frac{\partial}{\partial x}$ on (2):

$$\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}\int_{-\infty}^{+\infty} dp\, e^{ipx/\hbar}\frac{1}{\sqrt{2\pi\hbar}}\tilde{\psi}(p) = \int_{-\infty}^{+\infty} dp\, \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\, p\,\tilde{\psi}(p)$$

So, we have that

$$\frac{\hbar}{i}\frac{\partial}{\partial x}\psi \xleftrightarrow{\text{Fourier}} p\,\tilde{\psi}(p)$$

in the two spaces, $\hat{p}$ work in this way!
What we are saying is that, the operator $\hat{p}$ acts in different ways in the two representations!

$$\hat{p}\,\tilde{\psi}(p) \equiv p\,\tilde{\psi}(p) \quad !!!$$

Just like x!

[Io, spiegare: così usando le stesse formule e poi sfruttando le proprietà della trasformata di Fourier, ora da un punto di vista fisico è solo una diversa rappresentazione. Prima forza x in rappresentazione x, poi p in rappr. p, poi avevi le 2

So, in momentum space, $\hat{x}$ will be a derivative! Show it!

[SPIEGATO COSA È $\psi(x)$, DIRE CHE SI PUÒ SCRIVERE TRASFORMATA DI FOURIER COME $\tilde{\psi}(p)$, E DOPO parlare di MOMENTUM!]

We give data (ψ, $\tilde{\psi}$) in 2 different ways and so the operators to give the same effect, must operate in different ways!

time for a change of subject! :-)

Stern-Gerlach experiment and spin

We are going to work on spin-$\frac{1}{2}$ for most of these lectures.

Such experiment was done in Frankfurt, 1922. This was quite confusing. Pauli discovered that the electron has "two degrees of freedom", but what they were was not clear. Kronig suggested that they could have something to do with the rotation of the electron, but Pauli was not convinced: rotation should have been too fast! To have momentum! This could disintegrate.
Then Uhlenbeck and Goudsmit, in 1925, had this same "angular momentum" idea, and their adviser, Ehrenfest, said "this does not make so much sense but publish it". They are given credit for discovering the spin of the electron.
Then, 1927, after the experiment, everything was clear.
Stern and Gerlach wanted to measure speeds of thermal motion of ions. So, send beams of these ions and put magnetic fields, deflect them and measure velocity. As Bohr discussed the angular momentum of protons, they tried to detect it! But they tried with silver and in silver electrons have no orbital angular momentum but just spin!
Given the electron beam, it was SPLITTED: no one understood why not just deflected.

Space quantization => (ugly name)

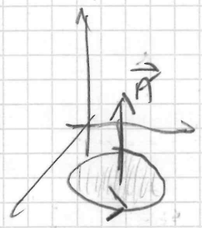
Let us study then this experiment and get some quantum mechanics lesson out of it.

First, important point. We can not "see" spin. We can "see", "measure", magnetic moments. Magnetic moments are identified by μ. They are the "magnetic analog" of electric dipoles: μ is a "magnetic dipole" moment.

The magnetic moment μ is given by

$\vec{\mu} = I \vec{A}$ → area; → current

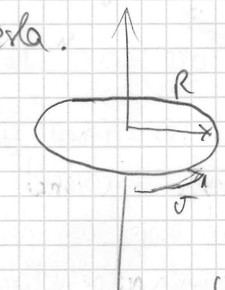
Roughly, we can think of the magnetic dipole as a loop in a plane, with an area vector $\vec{A}$ normal to it, direction determined by the right-hand rule by direction of current, flowing through the loop.



Units:

μB have units of energy. So, $[\mu] = \frac{J}{T}$ T is tesla.

In a charge configuration, we can assume to have a ring of charge, with radius R, total charge Q, and linear charge density λ. The ring rotates with speed v, has mass M



All this idea is aimed at introducing a relationship between spin, and magnetic moment (which is the thing, the quantity we can observe).

Current I is given by the linear charge density λ times velocity v:

$I = v\lambda$

But λ is the total charge divided by the lenght of the ring, $2\pi R$.

$\lambda = \frac{Q}{2\pi R}$

So,

$I = \frac{Qv}{2\pi R}$

The area $\vec{A}$ is πR^2, so

$\vec{\mu} = \frac{Q}{2\pi R} v \pi R^2 = \frac{1}{2} Q v R$

$\vec{\mu}$ is supposed to go up. Also $\vec{L}$, the angular momentum. In fact,

$L = M v R$ mass × radius × velocity: ANGULAR MOMENTUM!

So, we can observe, by multiplying and dividing μ (

$\mu = \frac{1}{2} Q v r = \frac{1}{2} \frac{Q}{M} M v R$,

that

$\mu = \frac{1}{2} \frac{Q}{M} L$!! μ and L are connected!

$$\boxed{\mu = \frac{Q}{2M} L}$$

This is very interesting as the "incidentals" such as velocity and radius dropped out! Just μ and L!

This is a universal relation! For every axisymmetric object: a sphere, or whatever!

This observation might lead us that angular momentum arises from rotation velocity, but this is just what Pauli didn't like.

What happens with quantum particles?

It: instead of writing L, we write S: spin angular momentum!?

For the electron, is it true that

$\mu = \frac{e}{2m_e} S$? e: electron charge; S: spin angular momentum; m_e: electron mass

Is this true? Not so much, but we can improve this formula! Let us multiply and divide times $\hbar$:

$\mu = \frac{e\hbar}{2m_e} \frac{S}{\hbar}$ ($\frac{e\hbar}{2m_e}$: units of magnetic moment; $\frac{S}{\hbar}$: this is dimensionless; in fact, $[\hbar] = [x\,p]$)

So, to correct it, we need a fudge factor: "g". Lande factor

$\mu = g \frac{e\hbar}{2m_e} \frac{S}{\hbar}$

sometimes g can be calculated, predicted

For an electron,

$g = 2$.

This is also predicted by Dirac, relativistic, equation. Yay!

The quantity

$\frac{e\hbar}{2m_e}$ is called μ_B: "Bohr magneton".

$\mu_B \approx 9.3 \times 10^{-24} \frac{J}{Te}$

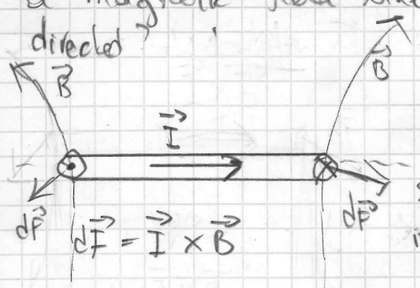
Notice: the electron, to the best of our knowledge, is a point particle. It has no spatial extension. Even though it has a mass, and an angular momentum. How can it has angular momentum an object with 0 radius? Well, it has. Period. =C

Since the electron has negative charge, e is -1.6×10^{-19} C, and

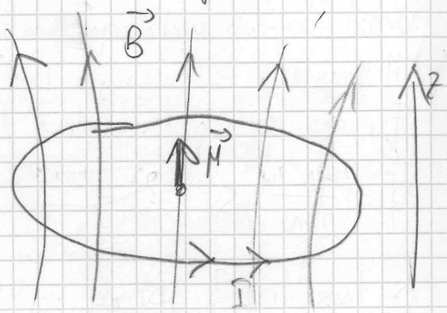
$\vec{\mu} = -2 \mu_B \frac{\vec{S}}{\hbar}$!!!

electron negative ch. — 2 — positive

So this is the story about what magnetic moments are.
Now. How do magnetic moments react to the presence of magnetic fields?
Well, if we have a loop of charges, and a magnetic field like this, where is force directed?



If the field lines are diverging, are curving, since $\vec{F}$ is normal to them, it has a component in ↓ direction (right-hand rule)



In fact,

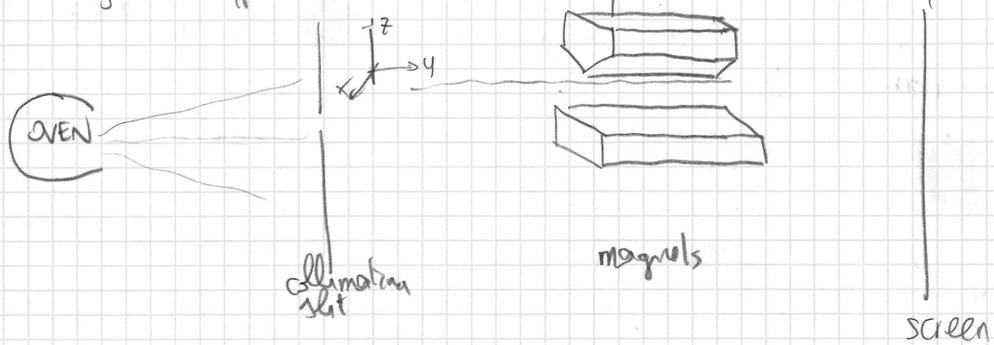
$\vec{F} = \nabla(\vec{\mu} \cdot \vec{B})$ (derived in electromagnetic courses)

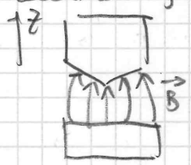
Since $\vec{\mu}$ goes up for hypothesis, so it is going in the same direction of $\vec{B}$.
As $\vec{B}$ is "more parallel" to $\vec{\mu}$ for $z > 0$ ($z = 0$ is the axis), as they didn't started diverging, $\vec{\mu} \cdot \vec{B}$ grows when going towards $-\vec{z}$: $\nabla(\vec{\mu} \cdot \vec{B})$ is NEGATIVE! (63min)
If the magnetic field is roughly directed along $\hat{z}$ we have simplifications.

Stern and Gerlach worked with silver atoms, which have 47 electrons.
46 fill up the levels, 1, 2, 3, 4, and a "5s" electron, which is a state with 0 angular momentum (s state), which is the one that enters in the apparatus. No electrons (almost) with angular momentum enter in the apparatus: just spin. (the unpaired electron)

Stern and Gerlach wanted to observe where electrons go when the associated dipoles reacted with the magnetic field

Stern-Gerlach apparatus: an oven produces silver atoms, that leave as a gas; then they reach a collimating slit, and then magnets, and a screen.



The magnets look like 

(so that the magnetic field has a gradient), lines bend

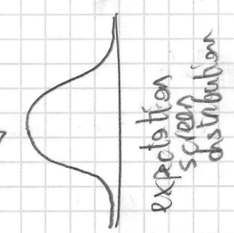
What did they get on the screen?

The force is basically in the z direction, as roughly $\vec{B}$ is parallel to $\vec{z}$,

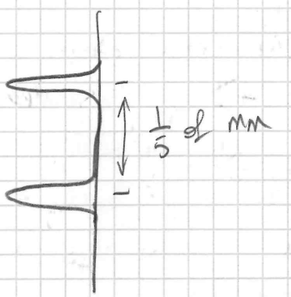
$\vec{F} = \nabla(\vec{\mu} \cdot \vec{B}) \approx \nabla(\mu_z B_z) = \mu_z \nabla B_z \approx \mu_z \frac{\partial B_z}{\partial z}$

The bending of the lines is negligible, as done in these steps.

As the oven generates electrons with random orientation, some deflected and so on, the expectation was something like ↗



Because the ones with many z component of the angular momentum get too deflected by $\vec{B}$ and don't illuminate the screen.
What was actually obtained was: 2 peaks! ↘



This is why they called it "space quantization"!

This confused everybody.
Then, after the story presented before, it was possible to understand

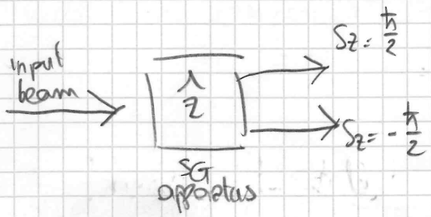
$\vec{\mu}_z = -2\,\mu_B \frac{S_z}{\hbar}$ which we wrote before.

Knowing B_z (that is decided during the experiment), it is possible to compute the deflection length, and it was possible to evaluate that

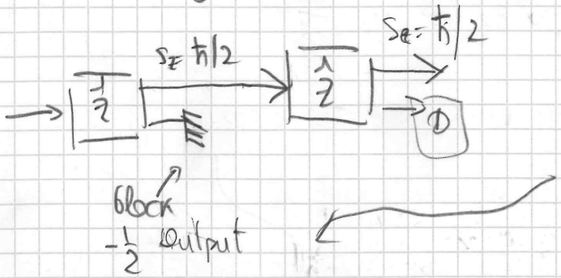
$\frac{S_z}{\hbar} = \pm\frac{1}{2}$!!!

So. Knowing that, mathematically, every state contains a superposition of spin-up $(+\frac{1}{2})$ and down $(-\frac{1}{2})$, the magnetic field changes the wavefunctions but still it can go in both deflected positions. Then, the detector measures them; what makes the wavefunction collapse into the top or bottom spot is the screen, the detector, which measures the state.

Let us draw schematically this in terms of boxes:



We can play with this and discover new things. Experiments show that,



We can: start with a first apparatus, have 2 outputs, but Block the $S_z = -\frac{\hbar}{2}$ output. Put a second apparatus in series to the $+\frac{\hbar}{2}$ output, and we will see that NO $-\frac{\hbar}{2}$ electrons are deleted.

This means that, with this trick, we can separate, identify, two kinds of states:

$|z,+\rangle$ states that, with respect to z direction, have spin $+\frac{\hbar}{2}$

$|z,-\rangle$ " " " " " " " " " " " " $-\frac{\hbar}{2}$

We can use them as a basis and even the states going towards x can be written as a superposition of them.
This system can be seen as 2 dimensional complex vector space! 2 vectors allow to generate all possible spin configurations!

In algebra, it means that the "minus" and "plus" states, $|z,+\rangle$ and $|z,-\rangle$ are orthogonal:

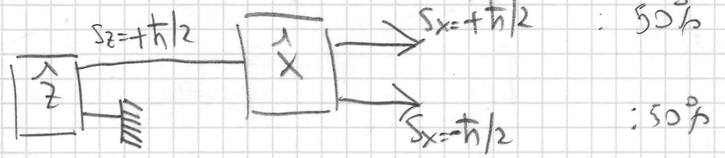
$\langle z-|z+\rangle = 0$!

Orthogonal means states! Even if spin vectors are both parallel, or antiparallel! NOTHING TO DO WITH IT, HERE!

Moreover,

$\langle z+|z+\rangle = 1$ every state that enter in the 2nd apparatus leave it!

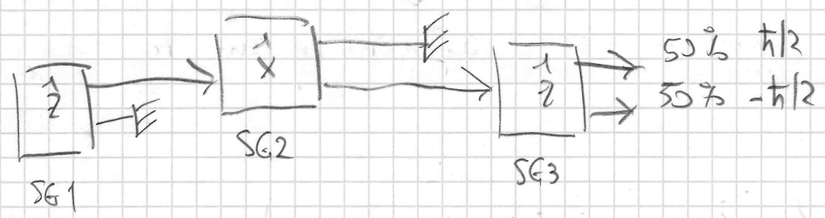
Another experiment : $\hat{z}$ SG apparatus and then $\hat{x}$ SG apparatus



half of the particles leave the $S_x = \hbar/2$ output, half the $-\hbar/2$ output!
So spin states along $\hat{x}$ have an overlap with those along $\hat{z}$!

$\langle x,+|z,+\rangle \neq 0$! OVERLAP!

Last experiment : $\hat{z}$, $\hat{x}$ and $\hat{z}$ again!



Both S_z

The relevant term is: every SG apparatus makes the state collapse into states parallel to their own direction: $\hat{z}$, $\hat{x}$ ... By this way as measuring make state collapse, in this experiment, even if we would expect 0% from the $S_z = -\hbar/2$ output, this is wrong, as SG2 changed the state!!!

Lecture 4: spin $\frac{1}{2}$; bras, kets, and operators

Last time we discussed the Stern-Gerlach (SG) experiment and its boxes, which can be put in cascade to understand the behavior of spin.
The key part in the stern-Gerlach machine was: a beam of silver atoms, which really are like electrons (they have no orbital angular momentum) are used. Beams were directed ONLY in the z-direction!
From this, we decided to treat spin $-\frac{1}{2}$ particles as a 2-dimensional complex vector space: this is the space of possible states of spin $-\frac{1}{2}$ particles.
Today we will set this framework.

The spin states of this silver atom could be described by the states

$|z;+\rangle$ and $|z;-\rangle$

$|z;+\rangle$: angular momentum $S_z = +\frac{\hbar}{2}$

$|z;-\rangle$: $S_z = -\frac{\hbar}{2}$

These are our 2 states.
The z label indicates that the atom passed a z-filter (Stern-Gerlach) state associated with S_z.

We could ask whether these states have some S_x or S_y component; we will discuss this in this lecture.

Mathematically, we state the fact that, for $|z;+\rangle$ to have $S_z = +\frac{\hbar}{2}$, means that there is an operator $\hat{S}_z$, which

$$\hat{S}_z |z;+\rangle = +\frac{\hbar}{2} |z;+\rangle$$

if $\hat{S}_z$ acts on the state, we obtain the same state, times $+\frac{\hbar}{2}$. It is the usual measurement procedure applied to $\hat{p}$ or $\hat{x}$, now, on spin states!

A measurement in quantum mechanics consists of operators applied to states. Similarly,

$$\hat{S}_z |z;-\rangle = -\frac{\hbar}{2} |z;-\rangle$$ the bar

When this happens $|z;\pm\rangle$ are eigenstates of the operator $\hat{S}_z$
Our hypothesis is that these two states, suffice: are capable of representing the whole state space! along ALL directions!
By just $|z,\pm\rangle$, all spin states, also along x and y, are built as linear combinations of these two, which are then a basis!

$|z;\pm\rangle$ are a basis of a 2D complex vector space.
So, given $|\psi\rangle$ any possible spin state,

$$|\psi\rangle = c_1 |z;+\rangle + c_2 |z;-\rangle \quad , \text{ where } c_1, c_2 \in \mathbb{C}.$$

The set of possible states $\{|\psi\rangle\}$ are the general vectors in a 2D complex vector space. Complex, because c_1, c_2 are complex, and 2D because the basis has dimension 2: it consists of 2 vectors!

These "vectors", $|z;\pm\rangle$, are called "kets". Kets are vectors.
Let us call $|z;+\rangle$ and $|z;-\rangle$ the "first" and "second" basis states, respectively. The vectors must not be complex: they are just vectors! Complex is in coefficients!

This idea, without providing numbers may sound too abstract. To give some numbers, we have to introduce a "representation": some ways of associating a state as a vector.
Being in our notation $|z;+\rangle$, we will write it as $|1\rangle$, and we will write $|z;-\rangle$ as $|2\rangle$.
If this is a 2D vector space, then, what is a vector? Well, from linear algebra courses, they are triplets or n-plets of numbers. So, to introduce a REPRESENTATION, we will say that

$$|z;+\rangle \triangleq |1\rangle \xrightarrow{\text{representation}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

we represent it as this vector!
A representation! A way of thinking about it

It is a way of associating the ket to a canonical basis vector of a 2D vector space!

Then, similarly,

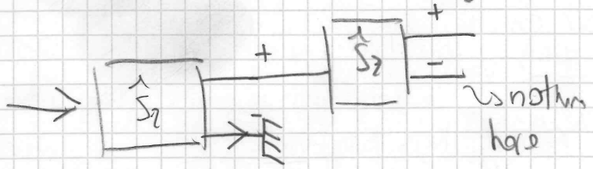
$$|z;-\rangle = |2\rangle \longleftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Therefore,

$$\psi = c_1 |z;+\rangle + c_2 |z;-\rangle \Leftrightarrow c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

we can represent $|\psi\rangle$ as a vector of 2 components!

The physical assumption that every spin state can be represented as a linear superposition of $|1\rangle$ and $|2\rangle$ is far from being clear, as of this moment. Nevertheless let us try to recover and exploit some of the ideas of the Stern-Gerlach experiment. In particular, we described an experiment with a $\hat{S}_z$ box, that filtered the $|z;+\rangle$ state, and put it in another $\hat{S}_z$ machine, and all the states went out through the top!



↳ nothing here

This means, physically, that a $|z;+\rangle$ state contains no amplitude of $|z;-\rangle$ states! orthogonal

Physically, these states are orthogonal!
This will require the whole framework, in detail, of bras and kets, to write this statement precisely; now we will just introduce few details, and later we will complete this!
This statement is that the overlap

$\langle z;-|z;+\rangle = 0$ this is the physical statement saying that all particles leave the top output!

We assume states to be normalized:
$\langle z;+|z;+\rangle = 1$!
If similarly we check the other part ...

$\langle z;+|z;-\rangle = 0$, $\langle z;-|z;-\rangle = 1$

[Notice [COMMENTS MIO]: seems like that bra is what we try to measure and ket is what we introduce in the SG box.
bra: output. ket: input.]

We've written 4 equations, but let's rewrite them with a more handy notation!

$\langle i|j\rangle = \delta_{ij}$ [in fact, $\langle 2|1\rangle = 0$, $\langle 1|1\rangle = 1$, ...]

We have not defined "bras", and we will do it for completeness now.
While we use $|1\rangle \longleftrightarrow \begin{bmatrix}1\\0\end{bmatrix}$ as representation for the first ket vector, we use $\langle 1| \longleftrightarrow [1 \;\; 0]$ for the first bra vector.

Similarly for $|2\rangle$ and $\langle 2|$ Bra: row vector ket: column vector!

$|1\rangle \longleftrightarrow \begin{bmatrix}1\\0\end{bmatrix}$; $\langle 1| \longleftrightarrow [0 \;\; 1]$ $|2\rangle \longleftrightarrow \begin{bmatrix}0\\1\end{bmatrix}$; $\langle 2| \longleftrightarrow [0 \;\; 1]$

Now, let's do this a little more generally. Given a state $|2\rangle =$

$|2\rangle = \alpha_1|1\rangle + \alpha_2|2\rangle, \longleftrightarrow \begin{bmatrix}\alpha_1\\\alpha_2\end{bmatrix}$

$|\beta\rangle = \beta_1|1\rangle + \beta_2|2\rangle \longleftrightarrow \begin{bmatrix}\beta_1\\\beta_2\end{bmatrix}$

The definition of a general bra, $\langle\alpha|$, is:

$\langle\alpha| = \alpha_1^* \langle 1| + \alpha_2^* \langle 2| \longleftrightarrow \alpha_1^* [1 \;\; 0] + \alpha_2^* [0 \;\; 1] = [\alpha_1^* \;\; \alpha_2^*]$

complex conjugate!

Ket: column vector representation of the state
bra: row vector representation of the state in which coefficients of the linear combination are the complex conjugates of those of ket

What is the bra-ket? It is a number; the reason for complex conjugating the bra is that

$\langle 2|2\rangle$ should return a positive number, like a length squared!

So, if we compute

$\langle\alpha|\beta\rangle$, we get a NUMBER, by exploiting the vector representation, using the matrix multiplication rules!

So

$\langle\alpha|\beta\rangle = \begin{bmatrix}\alpha_1^* & \alpha_2^*\end{bmatrix}\begin{bmatrix}\beta_1\\\beta_2\end{bmatrix} = \alpha_1^*\beta_1 + \alpha_2^*\beta_2$ this is the number called "inner product" or "bra-ket"

This is what is behind the formulas written before! And, then, vectors satisfying $\langle i|j\rangle = \delta_{ij}$, are called orthonormal: orthogonal and normal!
We discussed the representation of states, now, let us talk about how to represent the operator $\hat{S}_z$ performing the measurements!
We have to understand how $\hat{S}_z$ acts on these two component vectors! $|1\rangle$ and $|2\rangle$!

Firstly, it should be a 2×2 matrix, as it is the only object "acting" naturally on a 2 component vector (row & column!)

Claim: $\hat{S}_z$ can be represented by

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

we put "=" but "⟺" should be better: this is the representation of the operator!

This representation is a consequence of the fact that we started represented states as vectors!

So:

$\hat{S}_z|1\rangle \overset{?}{=} \frac{\hbar}{2}|1\rangle$? Yes!

$$\frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad ! \qquad \hat{S}_z|2\rangle \overset{?}{=} \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{\hbar}{2}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad !!$$

The operator does what we expect from it on the basis vectors, so, on! The representation is then ok ∀ states, as we will represent as combinations of basis vectors!

But, this is ok for $\hat{S}_z$. Now: what if we align the SG machine along $\hat{x}$, a $\hat{y}$, so that to measure S_x or S_y?

How can we know that, with just a $\hat{z}$ spin basis, we could also describe spin along x?! So just change the linear combination weights?

We have to invent something, to go on with our considerations.

Feynman had arguments based on rotating the Stern-Gerlach machines, but it's a little hard to follow. We will exploit some property of angular momentum: in fact, spin is a form of angular momentum as well!!

Another way of asking the questions about spin along x or y could be:

"what are the operators $\hat{S}_x$ and $\hat{S}_y$? Well, the analogy with angular momentum is useful, as we have $\hat{L}_z$, but $\hat{L}_x$ and $\hat{L}_y$ too!

From 8.04, we know something about of these operators!

We know that

$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$ — we know how to apply it to a wavefunction!

nd like $\hat{S}_z$, that requires the vector space framework and so on! This, see Allan Adams notes, because spin cannot be represented with wavefunctions!

We also know that $\hat{L}_z$ is Hermitian! We can then measure it! Well, $\hat{S}_z$ is Hermitian as well: if we transpose-complex conjugate the $\hat{S}_z$ matrix, we get the same matrix!

Recalling $\hat{L}_x$, $\hat{L}_y$, if we define $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ as

$\hat{L}_x = \hat{L}_1$, $\hat{L}_y = \hat{L}_2$, $\hat{L}_z = \hat{L}_3$,

ϵ_{ijk}: Levi-Civita tensor!

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k \qquad \epsilon_{123} = 1$$

this was called the "algebra of angular momentum"! We will get some practice on what this ϵ_{ijk} is!

For our purpose, let's write more explicit formulas:

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z,$$
$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x,$$
$$[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$$

this holds for orbital angular momentum.

Starting from here, trying to find a complete analogy: we're going to believe that $\hat{S}$ is similar, analog, to angular momentum. So,

$$[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z, \qquad [\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x, \qquad [\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y \qquad (1)$$

Can we get $\hat{S}_x$ and $\hat{S}_y$ from this analogy?

We know that these operators, so in this case these 3×3 matrices must be Hermitian to be related to observable quantities. So, let's focus on this point.

A Hermitian 2×2 matrix is:

$$\begin{bmatrix} 2c & a-ib \\ a+ib & 2d \end{bmatrix} \qquad (2)$$

where $c, d \in \mathbb{R}$
$a, b, \in \mathbb{R}$.

This is a Hermitian matrix: the complex conjugate transpose is equal.

The set of Hermitian matrices can be thought, formally, as a vector space.

Well. If $\hat{S}_x$ or $\hat{S}_y$ have something to do with identity matrices, they would commute with everything, and this would not be compatible with our physics, with our (1)!

So, let us remove, from (2), the matrix $(c+d)\mathbb{I}$

READ LECTURE NOTES!

$$\longrightarrow \begin{bmatrix} c-d & a-ib \\ a+ib & d-c \end{bmatrix}$$

this is still Hermitian. Let us try to find $\hat{S}_x$, $\hat{S}_y$ among these matrices. But. We already know $\hat{S}_z$. And, there is nothing bad if $\hat{S}_x$, $\hat{S}_y$ contain "a piece" of $\hat{S}_z$! To avoid confusing the situation, let us "disentangle" $\hat{S}_z$ from this matrix, subtracting it!

We can add to this matrix some multiple of $\hat{S}_z$ and kill the $c-d$ $d-c$ part, which is similar to $\hat{S}_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$!

↳ by adding some multiple of $\hat{S}_z$, we finally get

$$\begin{bmatrix} 0 & a-ib \\ a+ib & 0 \end{bmatrix}$$

we are restricting the form of the matrix suitable to represent $\hat{S}_x$, $\hat{S}_y$!

This matrix can be rewritten as:

$$= a\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

The first Hermitian matrix (times a) + the second (times b) is Hermitian: if we take a Hermitian matrix and multiply by a real number, the matrix remains Hermitian!

In some sense, then, the set of 2x2 Hermitian matrices is a real vector space, with 4 basis vectors! These are:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \text{ the } \hat{S}_z \text{ part } \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \text{ and the identity } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

4 basis "vectors" contributing to the general Hermitian matrix (2)!

These four 2x2 matrices are, in some sense, sort of linearly independent elements of a "vector space".

We are looking for $\hat{S}_x$ and $\hat{S}_y$, and we have two matrices which could be interesting for such job: $\hat{S}_z$ is not, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ rather, the remaining <u>two</u>!!

We are still missing information about the <u>scale</u> of $\hat{S}_x$, $\hat{S}_y$! Since the eigenvalues of $\hat{S}_x$ and $\hat{S}_y$ should be also $\pm\frac{\hbar}{2}$, just like $\hat{S}_z$!

Notice that (1) contains ambiguities: if we substitute $\hat{S}_x$ with $-\hat{S}_y$ and $\hat{S}_y$ with $+\hat{S}_x$, nothing changes!

So, there is no "unique" representation! Let us write the most popular one. Before that, let us turn to eigenvectors and eigenvalues.

$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has 2 eigenvalues: $\lambda_1 = 1$, with $\underline{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\frac{1}{\sqrt{2}}$,

$\lambda_2 = -1$ with $\underline{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}\frac{1}{\sqrt{2}}$

instead,

$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ $\lambda_1 = 1$, $\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ i \end{bmatrix}$ where normalization is: $\langle \underline{u}_i | \underline{u}_i \rangle = 1$

$\lambda_2 = -1$, $\underline{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$

With $\lambda_i = \pm 1$, it means that it should be sufficient to multiply our matrices times $\hbar/2$!

$$\hat{S}_x \stackrel{?}{=} \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \hat{S}_y \stackrel{?}{=} \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

Let us check (1):

$$[\hat{S}_x, \hat{S}_y] = \hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x = \left(\frac{\hbar}{2}\right)^2\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \left(\frac{\hbar}{2}\right)^2\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} =$$

$$= \left(\frac{\hbar}{2}\right)^2\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = \left(\frac{\hbar}{2}\right)^2\begin{bmatrix} 2i & 0 \\ 0 & -2i \end{bmatrix} = i\hbar\,\frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \to \hat{S}_z$$

$$= i\hbar\hat{S}_z$$

So, we have the 3 matrices! Yay! And, by construction, they are Hermitian! They have been defined as

$$\hat{S}_i = \frac{\hbar}{2}\sigma_i \quad \text{, where } \sigma_i \text{ is called "Pauli matrix"}$$

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Now. This gives us answers to most of the experiments we could do.

For example, now we have tools to see

$\hat{S}_x |x;\pm\rangle = \pm\frac{\hbar}{2}|x;\pm\rangle$ just like for $\hat{z}, \hat{S}_z$!

We just studied the eigenvalues of $\vec{S}_x$!

The eigenstate $\lambda_1 = \lambda_1$ so $+\frac{\hbar}{2}$, is

$$|x;+\rangle \longleftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\left(\begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}\right) \longleftrightarrow \frac{1}{\sqrt{2}}\left(|z;+\rangle + |z;-\rangle\right)$$

we used the correspondence, and shown that $|x;+\rangle$ can be written using the $|z;\pm\rangle$: we stated, we claimed, that z-states were sufficient! And, this is the proof! Then,

$$|x;-\rangle = \frac{1}{\sqrt{2}}\left(|z;+\rangle - |z;-\rangle\right) \longleftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$$

and so on! NICE! Every state can be written starting from the z states!

These formulas can be also inverted:

$|z;+\rangle = \frac{1}{\sqrt{2}}\left(|x;+\rangle + |x;-\rangle\right)$ if we measure from an $\hat{S}_x$ filter,

$|z;-\rangle = \frac{1}{\sqrt{2}}\left(|x;+\rangle - |x;-\rangle\right)$ we have 2 states along z!

So, this answers the question

$\langle x;+|z;+\rangle = \frac{1}{\sqrt{2}}$ $\quad$ $\langle x;-|z;+\rangle = \frac{1}{\sqrt{2}}$

this is consistent with the Stern-Gerlach experiment!

The y states can be constructed in the same way:

$\hat{S}_y |y;\pm\rangle = \pm\frac{\hbar}{2}|y;\pm\rangle$,

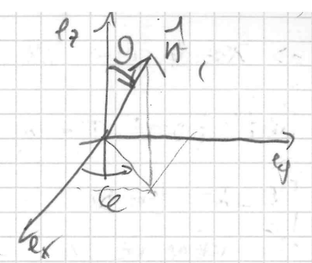
$|y;\pm\rangle = \frac{1}{\sqrt{2}}\left(|z;+\rangle \pm i|z;-\rangle\right)$; that's it!

Notice that we can build, from $|z\rangle$ states, 2 couples of orthogonal states ($|x\rangle$, $|y\rangle$) just thanks to complex numbers! ^^

Now, we have a theoretical framework that seems to describe the whole Stern Gerlach experiment!

Let us calculate states along an arbitrary direction: not just x, y, z!

$\vec{n} = n_x\vec{e}_x + n_y\vec{e}_y + n_z\vec{e}_z$, unit vector



how can we construct a spin state directed along $\vec{n}$?

Well.

$\vec{S} \to \begin{pmatrix}\vec{S}_x & \vec{S}_y & \vec{S}_z\end{pmatrix}$, so we can write it as

$\vec{S} = \hat{S}_x\vec{e}_x + \hat{S}_y\vec{e}_y + \hat{S}_z\vec{e}_z$! Where $\hat{S}_x, \hat{S}_y, \hat{S}_z$ are matrices times unit vectors! ^^

So, by using some intuition to write this "symbolic" relation,

$\hat{S}_{\vec{n}} = \vec{n}\cdot\vec{S} = n_x\hat{S}_x + n_y\hat{S}_y + n_z\hat{S}_z$

this is just an operator! Something "simple" to be rewritten! For example, if our interest is along z, $n_x = n_y = 0$! And so on!

$\hat{S}_{\vec{n}}$ is the spin operator in direction $\vec{n}$.

Let us study it. The polar coordinates of $\vec{n}$ are

$n_x = \cos\varphi \sin\theta$

$n_y = \sin\varphi \sin\theta$

$n_z = \cos\theta$

So,

$$\hat{S}_{\vec{n}} = \frac{\hbar}{2}\left(n_x\sigma_1 + n_y\sigma_2 + n_z\sigma_3\right) = \quad (n_x - in_y = \sin\theta(\cos\theta - i\sin\theta))$$

$$\frac{\hbar}{2}\begin{bmatrix}n_z & n_x - in_y\\ n_x + in_y & -n_z\end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix}\cos\theta & \sin\theta e^{-i\varphi}\\ \sin\theta e^{+i\varphi} & -\cos\theta\end{bmatrix} = \frac{1}{\hbar}\hat{S}_{\vec{n}}!$$

Which are the eigenvectors and eigenvalues?

$$\det\left(\hat{A} - \lambda\hat{I}\right) = 0 \Rightarrow \det\left\{\begin{bmatrix}\frac{\hbar}{2}\cos\theta - \lambda & \sin\theta e^{-i\varphi}\\ \sin\theta e^{+i\varphi} & -\frac{\hbar}{2}\cos\theta - \lambda\end{bmatrix}\right\} = 0?$$

= not so terrible.

$\Rightarrow \lambda = \pm \frac{\hbar}{2}$

what about the eigenvectors? We should solve the homogeneous system

$\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) \underline{v} = 0$

Let us look for $|\vec{n}; +\rangle$ for example. This satisfies

$\hat{S}_{\vec{n}} |\vec{n}; \pm\rangle = \pm \frac{\hbar}{2} |\vec{n}, \pm\rangle$.

This state can be written as a combination of z states:

$|\vec{n}; +\rangle = c_1 |z; +\rangle + c_2 |z; -\rangle \doteq \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

So,

$\left(\hat{S}_{\vec{n}} - \frac{\hbar}{2} \underline{\underline{I}}\right) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0$

$\hookrightarrow \begin{bmatrix} \cos\theta - 1 & \sin\theta e^{-i\varphi} \\ \sin\theta e^{+i\varphi} & -\cos\theta - 1 \end{bmatrix} \frac{\hbar}{2} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0$

actually, the two equations are linearly dependent, since we used an eigenvalue!

So,

$c_2 = e^{i\varphi} \frac{1-\cos\theta}{\sin\theta} c_1$,

but, we can simplify this, as it equals

$c_2 = e^{i\varphi} \tan\left(\frac{\theta}{2}\right) c_1$, and, normalized,

$\hookrightarrow |c_1|^2 \left(1 + \frac{\sin^2\theta/2}{\cos^2\theta/2}\right) = 1 \Rightarrow |c_1|^2 = \cos^2\frac{\theta}{2}$

So, I can choose

$c_1 = \cos\frac{\theta}{2}$ $\quad c_2 = \sin\frac{\theta}{2} e^{i\varphi}$ is a possible choice!

So,

$|\vec{n}; +\rangle = \cos\frac{\theta}{2} |z; +\rangle + e^{i\varphi} \sin\frac{\theta}{2} |z; -\rangle$

this is an amazing result!

Thus spin state can be represented with $|z; \pm\rangle$ states!

It can be shown that the $-$ state is:

$|\vec{n}; -\rangle = \sin\frac{\theta}{2} |z; +\rangle - \cos\frac{\theta}{2} e^{i\varphi} |z; -\rangle$

Notice that this has, for $\theta = 0$, the φ ambiguity, so it is convenient to multiply everything times $-e^{-i\varphi}$:

$\hookrightarrow |\vec{n}; -\rangle = -e^{-i\varphi} \sin\frac{\theta}{2} |z; +\rangle + \cos\frac{\theta}{2} |z; -\rangle$,

which solves the $\theta = 0$ problem :-)

Lecture 5: Linear Algebra: Vector Spaces and Operators

A note on Pauli matrices. We could notice that

$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathbb{1}$

This tells something interesting about the eigenvalues. Let us consider the following matrix equation:

$\underline{\underline{M}}^2 + \alpha \underline{\underline{M}} + \beta \mathbb{1} = \underline{\underline{0}}$, and assume that $\underline{\underline{M}}$ satisfies such an equation (1)

Let's say that we want to find the eigenvalues of $\underline{\underline{M}}$:

$\underline{\underline{M}} \underline{v} = \lambda \underline{v}$

Then, let act the matrix on the left-hand side of (1) act on the eigenvector $\underline{v}$:

$\underline{\underline{M}}^2 \underline{v} + \alpha \underline{\underline{M}} \underline{v} + \beta \underline{v} = 0$

$\hookrightarrow \lambda^2 \underline{v} + \alpha \lambda \underline{v} + \beta \underline{v} = 0 \iff \boxed{\lambda^2 + \alpha\lambda + \beta = 0}$ for nontrivial $\underline{v}$! an eigenvector is not 0! Otherwise λ would be not determined!

So, as $\sigma_1^2 = \mathbb{1}$,

$(\lambda)^2 = 1$, then $\lambda_{1,2}$ could be $1, \pm 1, -1, \dots$

Another important thing is the trace of the matrix!

$Tr\,\sigma_1$ = sum of elements on diagonal = (0)

Actually, the trace of all Pauli matrices appear to be 0.
A theorem of linear algebra shows that the trace of a matrix is equal to the sum of its eigenvalues!
So, for Pauli matrices, if the eigenvalues can only be $+1$ or -1, then due to the "trace" theorem if one eigenvalue is $+1$, then the other is -1.
The lecture notes discuss these points, as well as other properties of Pauli matrices. Two basic properties are, recalling

$$\hat{S}_i = \frac{\hbar}{2}\sigma_i,$$

and from the algebra of angular momentum, $[\hat{S}_i, \hat{S}_j] = \epsilon_{ijk}\hat{S}_k$,

$$\hookrightarrow [\sigma_i, \sigma_j] = 2i\,\epsilon_{ijk}\,\sigma_k$$

This, concerning commutators. But, Pauli matrices exhibit interesting properties also dealing with anticommutators! Even if, as appears, σ_1 and σ_2 don't commute, they satisfy some other property:

$\sigma_1\sigma_2 = -\sigma_2\sigma_1$: anticommutation!

$$\hookrightarrow \sigma_1\sigma_2 + \sigma_2\sigma_1 = 0$$

$\hookrightarrow \{\sigma_1, \sigma_2\}$: anticommutator of σ_1 and σ_2.

$$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$$

A little more analysis in the lecture notes show that

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\,\mathbb{1}$$

With this result, we get a general formula: any product of two operators, $\hat{A}\hat{B}$, can be written as

$$\hat{A}\hat{B} = \frac{1}{2}[\hat{A}, \hat{B}] + \frac{1}{2}\{\hat{A}, \hat{B}\} \qquad \text{! Nice!}$$

So,

$$\sigma_i\sigma_j = \delta_{ij}\,\mathbb{1} + i\epsilon_{ijk}\sigma_k \qquad (2)$$

In order these formulas to be neater, let us invent a notation in which we think of $\vec{\sigma}$ as a triplet:

$$\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3),$$

$\vec{a} = (a_1, a_2, a_3)$, and the product of $\vec{a}$ (vector) times $\vec{\sigma}$ is:

$$\vec{a}\cdot\vec{\sigma} \triangleq a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \;!$$

$$= a_i\sigma_i \quad (\text{summing over } i\,!)$$

this allows to interpret (2) nicely! Let us compute

$a_i\sigma_i b_j\sigma_j$, intended as $(\vec{a}\cdot\sigma)(\vec{b}\cdot\sigma)$. Using the relation (2), this is equal to:

Sum over both i and j

Quick notation!

$$a_i\sigma_i\,b_j\sigma_j = a_i b_j\,\sigma_i\sigma_j = a_i b_j\,\delta_{ij}\,\mathbb{1} + i\,\epsilon_{ijk}\,a_i b_j\sigma_k =$$

$$(\vec{a}\cdot\sigma)(\vec{b}\cdot\sigma) = (a_i b_i)\,\mathbb{1} + \cdots$$

$$= (\vec{a}\cdot\vec{b})\,\mathbb{1} + \cdots$$

How can we compute $\epsilon_{ijk}a_i b_j\sigma_k$? Well, k is a free index, so this is the k component of $\vec{a}\times\vec{b}$:

$$\vec{a}\times\vec{b} = \begin{bmatrix} \cdot & \cdot & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \quad a_1 b_2 - a_2 b_1$$

$$= (\vec{a}\times\vec{b})_k,$$ so that this is equal to

$$\hookrightarrow (\vec{a}\cdot\sigma)(\vec{b}\cdot\sigma) = (\vec{a}\cdot\vec{b})\,\mathbb{1} + i(\vec{a}\times\vec{b})\cdot\vec{\sigma},$$

as only the k component is involved and σ_x, σ_y don't contribute.

For instance, if $\vec{a} = \vec{b} = \vec{n}$, this is:

$$(\vec{n}\cdot\vec{\sigma})^2 = \mathbb{1} \qquad \text{as } \hat{n}\times\hat{n} = 0.$$

This identity allows to better understand $\hat{S}_{\vec{n}}$, which was

$$\hat{S}_{\vec{n}} = \vec{n}\cdot\hat{\vec{S}} = \frac{\hbar}{2}\vec{n}\cdot\vec{\sigma},$$

and if we square this,

$$\left(\hat{S}_{\vec{n}}\right)^2 = \left(\frac{\hbar}{2}\right)^2\mathbb{1}\;!$$

Moreover, the trace is 0 as it is the trace of a sum of matrices, all with 0 trace!!!

$Tr\left\{\hat{\vec{S}}_{\vec{n}}\right\} 0$,

so that the eigenvalues are $\pm\frac{\hbar}{2}$! They are $\sqrt{\left(\frac{\hbar}{2}\right)^2}$, and their sum is 0!
So, even in an arbitrary direction, the eigenvalues are always $\pm\frac{\hbar}{2}$!
There's a little map of an aside in here in the notes.
Let us assume to have two triplets of operators, $\hat{\vec{X}} = (\hat{X}_1, \hat{X}_2, \hat{X}_3)$, $\hat{\vec{Y}} = (\hat{Y}_1, \hat{Y}_2, \hat{Y}_3)$,

$$\hat{\vec{X}} \cdot \hat{\vec{Y}} = \hat{X}_i \hat{Y}_i \qquad \left(=0, \sum_i \hat{X}_i \hat{Y}_i\right)$$

In general, this may not commute: if $\hat{X}_i, \hat{Y}_i$ don't!

NOT COMMUTATIVE!
$\hat{\vec{X}} \cdot \hat{\vec{Y}} \neq \hat{\vec{Y}} \cdot \hat{\vec{X}}$ IN GENERAL!

Similarly, we can define the cross product

$$\left(\hat{\vec{X}} \times \hat{\vec{Y}}\right)_k = \epsilon_{ijk} \hat{X}_i \hat{Y}_j \qquad \left[= \sum_i \sum_j \epsilon_{ijk} \hat{X}_i \hat{Y}_j\right)$$

In general, the cross product is antisymmetric:

$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A}$, $\vec{A}, \vec{B}$ vectors

But here we have operators! That could not commute!
Even $\hat{\vec{X}} \times \hat{\vec{X}}$ could be $\neq 0$!
So, question, what is

$\hat{\vec{S}} \times \hat{\vec{S}}$? $= i\hbar\hat{\vec{S}}$

COURSE NOTATION: WHEN WE WRITE REPEATED INDICES, LIKE $\hat{X}_i \hat{Y}_i$, THIS IS IMPLICITLY $\sum_i \hat{X}_i \hat{Y}_i$, UNLESS SPECIFIED DIFFERENTLY!

24, nov

Linear Algebra Review

This could be useful for several reasons. For example, many physicists think that the matrix representation of an operator must use bra/ket; this is not true. Complex vector spaces are very different from real ones.

Vector spaces and dimensionality

We care about this because we understood that the states of our physical system can be treated just like vectors in a complex vector space.

Observables, moreover, are operators acting on such vectors.
It is useful to understand the properties of complex vector spaces, compare them to those of real vector spaces, and so on.
In a vector space, we have vectors, and numbers. If numbers are real or complex, the vector space is real or complex; vectors are always real.
Vectors can be added, and numbers multiplied by vectors, to achieve vectors.
$\mathbb{R}$ and $\mathbb{C}$ are fields: real and complex fields, for numbers.

A vector space V is a set of vectors with an addition operation (+) that assigns a vector $u+v$ in the vector space, when $u \in V$, $v \in V$.
[This means that a vector space is closed under addition: we cannot go "out" of V by adding vectors belonging to V.]

and a scalar multiplication by elements of the field $\mathbb{F}$ such as, for
$a \in \mathbb{F}$, $a v \in V$
[the space is also closed under multiplication by a scalar].

These operations satisfy the following properties.

1. $u+v = v+u$, $u, v \in V$ (commutative)
2. $u+(v+w) = (u+v)+w$, $u, v, w \in V$
   $(ab)v = a(bv)$, $v \in V$, $a, b \in \mathbb{F}$. (associative)
3. There is an additive identity: $\exists\, 0 \in V$, such that (additive identity)
   $v + 0 = v$, $v \in V$.
4. $\exists$ the number $1 \in \mathbb{F}$ such that
   $1 v = v$, $v \in V$
5. For each $v \in V$, there is $u \in V$ such that
   $u + v = 0$ for every element, you can find its opposite!
6. $a(u+v) = au + av$, (Distributivity)
   $(a+b)v = av + bv$, $a, b \in \mathbb{F}$, $v \in V$

A lot of definitions, but quite minimal!

Several interesting results can be proved starting from these elementary properties!
Considering for instance the additive identity, involving $\underline{0}$, it is easy to prove that this vector $\underline{0}$ is unique: if you find another vector, $\underline{0}'$, such that $\underline{v}+\underline{0}'=\underline{v}$, then $\underline{0}'=\underline{0}$!
You can also show that $0\underline{v}=\underline{0}$! Where $0 \in \mathbb{F}$, $\underline{0} \in V$!
These are 1-line exercises, all done in Axler book!
$2\underline{0}=\underline{0}$, and so on!
Moreover, you could prove that the additive inverse is unique, and it is called "$-\underline{v}$", and you can show it equals the number -1, times the vector $\underline{v}$!

Examples
Let us remark that vectors are NOT real or complex: vectors are vectors, and real/complex are just the numbers multiplying them.

Example (1): the set of N-component vectors:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}, \quad a_i \in \mathbb{R}, \quad i = 1 \dots N$$

[sum of 2 vectors: sum each component!]

This is a "vector space over $\mathbb{R}$", or "real vector space".

Example (2)
The set of $M \times N$ matrices with complex entries

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \dots & a_{1N} \\ a_{21} & & & \\ \vdots & & & \\ a_{M1} & a_{M2} \dots & & a_{MN} \end{bmatrix}, \quad \text{with } a_{ij} \in \mathbb{C},$$

is a complex vector space!

Example (3)
This is a little more surprising.
The space of 2×2 Hermitian matrices is a real vector space.
This is surprising because the entries of Hermitian matrices could be complex!

Like, as we said, $\begin{bmatrix} c+d & a+ib \\ a-ib & c-d \end{bmatrix}$! complex entries!
But, this is a real vector space because, if we multiply each element of the space, i.e., each Hermitian matrix, by a number, then the matrix remains Hermitian! If we multiply times a complex number we lose the Hermiticity, so that we "leave" the field! The space of Hermitian matrices is real, NOT complex!
Multiplication by real numbers preserves Hermiticity!
What does matter is NOT the fact that entries are real or complex, but the fact that multiply times real/complex numbers preserves or not the property of the element!

Example (4)
A set of polynomials $p(z)$, where $z \in \mathbb{F}$, and $p(z) \in \mathbb{F}$, where
$p(z) = a_0 + a_1 z + a_2 z^2 + \dots a_n z^n$, with n finite, $\{a_i\} \in \mathbb{F}$,
this is a vector space; $\mathcal{P}(\mathbb{F})$, space of all polynomials, is a vector space over $\mathbb{F}$.
Just like for matrices, "summing" polynomials means summing the coefficient of the corresponding powers of z.
Remark: n is finite, but still, this is ∞ dimensional: there are ∞ possible polynomials!
Example (5)
The set $\mathbb{F}^\infty$ of infinite sequences $(x_1, x_2, \dots)$, where $\{x_i\} \in \mathbb{F}$, is a vector space! It's like dealing with ∞-component column vectors!
The set of ∞ sequences is ∞! But still, it is a vector space over $\mathbb{F}$!

Example (6)
The set of complex functions on an interval $x \in [0, L]$, $f(x)$, is a complex vector space, with ∞ dimensionality

We mentioned the concept of dimensionality, but we haven't defined it rigorously. To this aim, we need to define the subspace of a vector space.

A subspace of a vector space is a subset of it, which is still a vector space itself. In particular, it is a subset that must contain the vector $\underline{0}$, and satisfy two properties.
A vector space, given this definition, can be represented as a sum of smaller vector spaces; this is useful to study, for instance, angular momentum.

Direct sum $\oplus$: we can write the vector space V as

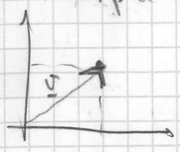
$$V = U_1 \oplus U_2 \oplus U_3 \oplus \dots \oplus U_m .$$

When we write this, we say that $\{U_i\}$ are subspaces of V, and any $\underline{v}$ in the vector space can be written uniquely as:

$$\underline{v} = a_1 \underline{u}_1 + \dots \underline{u}_2 + \dots + a_m \underline{u}_m, \quad \text{with}$$

$$\underline{u}_i \in U_i \quad !$$

In other words, we are decomposing our vector space V in some "basic ingredients"; this is behind the concept of direct sum.
For this to be true, $\underline{v}$ must be written as a sum of vectors belonging to each subspace! UNIQUELY! If the sum can be done in more than one way, then it means that the subspace overlap, and this is no longer a direct sum! The direct sum is applied to the decomposition into minimal terms!
For example $\mathbb{R}^2$: any vector $\underline{v}$ is a combination, UNIQUE, of



the two components! x, y!

So, $\mathbb{R}^2 = \mathbb{R}^1 \oplus \mathbb{R}^1$!

This idea of direct sum allows us to get the concept of dimensionality in a precise way.
Intuitively, the dimension is the number of linearly independent vectors we need to describe the whole set of vectors. But this view is not so useful for ∞ dimensions =/
Let us consider a "list of vectors", $(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n)$.
Any list of vectors has finite length n, by definition.
Given such list, which is its span? That is, how much do we reach by combining elements of that list? This is the span!

$$\text{span}(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n) = \{ a_1 \underline{v}_1 + a_2 \underline{v}_2 + \dots + a_n \underline{v}_n, \; a_i \in \mathbb{F} \}$$

The span of the list is the set of all possible products of vectors of the list times scalars belonging to the field $\mathbb{F}$, added (linear combinations)!
The span of the list "fills", "explores", the whole vector space!

So, the space V is finite-dimensional if it's spanned by some list. (because the list is finite-dimensional by definition)
And, well, it is infinite-dimensional, if it's not finite-dimensional ;-)
So, no list spans the space.
For instance, let's think about polynomials. If the space of polynomials was finite, then we would have a finite list of polynomials spanning the whole polynomial space. The point is that it is not the case: since the list is finite, then we have a polynomial of maximum degree, d_{max}. Well: a polynomial of degree $d_{max}+1$ is still a polynomial, but it is not in the span! $\Rightarrow$ No finite list can span the space of polynomials!
This explains more rigorously why example 4 was ∞-dimensional

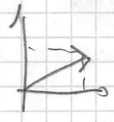
Example 1 was finite-dimensional! To represent a vector with N components, we need a list with N elements. For example, $N=3$,

$$\left[\begin{matrix}1\\0\\0\end{matrix}\right], \left[\begin{matrix}0\\1\\0\end{matrix}\right], \left[\begin{matrix}0\\0\\1\end{matrix}\right] \quad \text{is a list!}$$

A list $(\underline{v}_1, \dots \underline{v}_N)$ is linearly independent if

$$a_1 \underline{v}_1 + a_2 \underline{v}_2 + \dots a_N \underline{v}_N = \underline{0} \quad \text{has the unique solution } \{a_i\} = 0 \;\; \forall i !$$

This means that all coefficients MUST be set to 0 to obtain $\underline{0}$! This is the case of $\mathbb{R}^2$! The only way to achieve the null vector ("the origin") is $x=y=0$! =)



Now we can define a basis.
A basis of V is a list of vectors in V that spans V and it's linearly independent.
So, is a list with enough vectors to span the whole space, but not too much to avoid linear dependencies! It should be minimal!

Every finite-dimensional space has a basis!
Notice that the basis is not unique! We will change basis all the time, but a basis has always the same number of vectors! This length, is the DIMENSION of the vector space! The length of any basis of it!

Notice that it does not matter whether vectors are orthogonal or not. Defining orthogonality requires defining an inner product, and we have not mentioned it, because it is NOT NECESSARY for a set to have an inner product, to be a vector space. Inner product is something more. This is what we meant when we said that the matrix representation of an operator does not require bra-ket: the basis is a concept that does not require the concept of inner product.
The trace of an operator don't need the definition of an inner product! Neither eigenvalues nor eigenvectors! Inner product is needed just to get NUMBERS!
Before this, it is useful to show what we can obtain without introducing concepts!
Then, for example 1, since $\left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right)$ are linearly independent, then they are a basis, and, since the last has 3 elements, the dimension is $N=3$! =)
For example 2 it can be shown that it has dimension $N \times N$.
For example 3, Hermitian matrices (2×2), 4 elements!

$\mathbb{I}$, σ_1, σ_2, σ_3 are the 4 elements!
All of them are Hermitian, they span the space (we can just put arbitrary real numbers and obtain the general matrix)

$$a\sigma_1 + c\mathbb{I} + b\sigma_2 + d\sigma_3$$

the last is independent: there is no way to obtain all 0 coefficients without setting a, b, c, d to 0!
So, dimension 4!

To conclude this lecture, we will introduce a definition
Linear operator
When we talk about linear maps, we have two vector spaces V and W, and, in general, a map from V to W is something transforming an element of V into one of W.
In all generality, V and W may not have the same dimension!
When we have a linear map from V to V itself, this is called "linear operator".
A linear operator is a function $T : V \to V$ (T which turns V to V),

such that

$T(\underline{u}+\underline{v}) = T\underline{u} + T\underline{v}$, T acting on $\underline{u}+\underline{v}$ equals T acting on $\underline{u}$ + T acting on $\underline{v}$.

$T(a\underline{v}) = aT\underline{v}$

these two properties are "linearity"!
So, to know what a certain linear operator is, we just have to know how it acts on basis vectors! (because then any vector can be written as a linear combination of basis vectors, as a superposition!)
This will help us finding the matrix representation of an operator. This because

$T_{ij} = \langle i|T|j\rangle$, but, we don't need bras and kets to define T_{ij}: just, to COMPUTE it!

Lecture 6 : Linear Algebra : Vector Spaces and Operators (continued)

Aram Harrow

Last lecture we studied vector spaces from a more abstract, but also general, perspective, than just of sets of lists of numbers.
Today we will discuss operators. These, somehow, are related to matrices, applied to vectors, to transform them. The concept of operator can be applied also to ∞-dimensional spaces, though.
Moreover, today we will discuss also basis-independent quantities!

Which is the motivation of all this math, this abstract perspective? Well, operators are how we define observables; by this way, if we are interested in the properties of observables, such as Hamiltonians (energy), we need to study the properties of operators.
It is also important to discuss states; currently, states are defined as elements of a vector space. But we will also study, in the future, a formalism in which states are described as "density operators", or "density matrices". Moreover, operators are very important to study symmetries in quantum systems!

In the last lecture we defined a linear map $\mathcal{L}(V, W)$ from the

vector space V to W; linear means that, for all $\underline{u}, \underline{v} \in V$,

$T(\underline{u}+\underline{v}) = T\underline{u} + T\underline{v}$, $T(a\underline{v}) = aT\underline{v}$, $a \in \mathbb{F}$.

As we already stated, a relevant case is $\mathcal{L}(V,V)$: from V to V itself. They are called "operators on V".
Let us consider few examples.

(1) For vector space V, let us consider $\mathbb{R}(x)$ the vector space of all real polynomials in 1 variable $x \in \mathbb{R}$.
We can define various operators on such space. For example, differentiation!

$$T = \frac{d}{dx} \longrightarrow Tp = p'$$

this is a "function", a "map", which transforms polynomials into polynomials and it is also linear (linearity of derivative ;-)
Another operator, S, is the multiplication by x:

$S = \hat{x} \longrightarrow Sp = xp$ multiply polynomial by x. This is again linear, and returns another polynomial.

Even if we are talking about operators and we claimed some correspondence with matrices, well, it is clear that the idea is much more general ;-)

(2) $V = \mathbb{F}^\infty$: vector space of ∞ sequences!

$$= \{(x_1, x_2, \dots) : x_i \in \mathbb{F}\}$$

One operator we can define is L : left-shift operator!
Shift the entire sequence "left" and remove the first entry

$$L(x_1, x_2, x_3 \dots) = (x_2, x_3, x_4 \dots)$$

Again, we have a sequence, and it is linear
Right-shift, the same:

$$R(x_1, x_2, x_3 \dots) = (0, x_1, x_2, x_3 \dots)$$

R creates some room for a new number, that we set to 0 as the only choice, otherwise the new sequence would not be linear =(.
Otherwise, we could map a $\underline{0}$ vector into a non-zero vector

By linearity that's impossible, as we could have any vector, multiply it by the scalar 0, and it should equal to 0:

$$R(0(x_1, x_2, x_3, x_4 \dots)) \overset{?}{=} 0\, R(x_1, x_2, x_3 \dots) \quad ?$$

Not if I don't set to 0 the first slot!

$$R(0(x_1, x_2, x_3 \dots)) = R(0x_1, 0x_2, 0x_3 \dots) = R(\overset{?}{0}, 0x_1, 0x_2, 0x_3)$$

which must equal $0(?, x_1, x_2, x_3 \dots)$

(3) Another example of operator is the $\hat{0}$ operator,

$\hat{0}\,\underline{v} = \underline{0}$,

(4) $\hat{\mathbb{I}}$: identity operator

$\hat{\mathbb{I}}\,\underline{v} = \underline{v}$ sends the vector to itself

The space $\mathcal{L}(V)$ of all possible operators on V has some properties, which we are going to discuss. This is the set of all linear maps from V to itself, a set of maps, but this is also a vector space by itself!
So, the set of operators satisfies all properties of a vector space! It contains a $\hat{0}$ operator, closed under addition of vectors, under scalar multiplication, etc. So, everything we can do on every other vector space, such as finding a basis and so on, we can do it here as well.
But, in addition of having a vector space structure (closeness with respect to addition and multiplication times scalar), has "multiplication", which means, we could multiply operators and still obtain other operators.
This is a peculiarity of operators, not of generic linear maps, because the fact that applying multiplication returns another operator is a consequence of the fact that it maps elements of V to V, not to W!
This is like dealing with square matrices: with rectangular ones, given $\underline{\underline{A}}, \underline{\underline{A}}' \in \mathbb{R}^{m,n}$, $\underline{\underline{A}}\,\underline{\underline{A}}'$ is not defined, due to mismatch of dimensions.
given $\underline{\underline{B}}, \underline{\underline{B}}' \in \mathbb{R}^{n,n}$, $\underline{\underline{B}}\,\underline{\underline{B}}'$ exists.

The existence of multiplication does not mean that it has all the properties of the "standard" multiplication in fields.

It has associative property, i.e.,

$A(BC) = (AB)C$,

and to prove this, we can check how these two act on vectors. Indeed, the behavior of an operator is characterized by how it behaves applied to vectors! So, how does act (AB) on an arbitrary vector $\underline{v}$?

$(AB)\underline{v} \triangleq A(B\underline{v})$: first apply B on $\underline{v}$. Then, A on what results.

This is the definition. And proves associativity.

Another property that we have in operators: identity ($\exists\ \hat{\mathbb{I}}$!)

But, there are other properties we don't have; in fact, inverses are not always defined. Think of matrices: the inverse is not always defined!

Another lacking property is: operators or, matrices, are not necessarily commutative! For instance, for polynomials,

$$T\,p = x\frac{d}{dx}p \quad ; \qquad TS\,x = \frac{d}{dx}\left[x\,p\right]$$

and these two are clearly different!

Often, in quantum mechanics, we compute to which extent they do not commute; this leads to defining the commutator

$[A,B] = AB - BA$,

and, for instance, considering in our example $p = x^n$,

$$ST\,x^n = x\frac{d}{dx}x^n = x\,(n)\,x^{n-1} = n x^n$$

$$TS\,x^n = \frac{d}{dx}\left[x\,x^n\right] = \frac{d}{dx}x^{n+1} = (n+1)\,x^n,$$

and then the commutator is

$$\left[T,S\right]x^n = (TS - ST)x^n = \frac{d}{dx}\left[x\,p\right] - x\frac{dp}{dx} = p + x\frac{dp}{dx} - x\frac{dp}{dx} = p,$$

which is, in our case, x^n! So,

$$\left[T,S\right]p = p = \hat{\mathbb{I}}\,p.$$

the commutator of these two operators is the identity!

We discussed the properties of the space of operators. Now, we are going to discuss the properties of individual operators.

The gross features of an operator could be got by looking at some of its mathematical properties. One is its nullspace, or "kernel":

$$\text{Ker}\,T = \left\{ \underline{v} \in V \ : \ T\underline{v} = \underline{0} \right\} \supseteq \underline{0}$$

it is the set of vectors such that, when we apply the operator T to them, go to $\underline{0}$! The nullspace always includes the null vector $\underline{0}$, but it could also include other elements. For instance, $\hat{\mathbb{I}}$ has only the vector $\underline{0}$.

The nullspace of an operator is a subspace of V, as the sum of two vectors belonging to it is mapped to $\underline{0}$ by the operator, as well as a vector times a scalar.

The fact that the nullspace is a subspace of V is useful to understand some of its properties. Let us try to think of T as a function (it is a mapping). With functions, we try to understand if they are injective, bijective, surjective. It is known for functions that T is injective if, for $\underline{u} \neq \underline{v}$, $T\underline{u} \neq T\underline{v}$: T maps distinct vectors to distinct vectors. It turns out to be equivalent to the null space being only the $\underline{0}$ vector.

Another way of saying this is, for $\underline{u} \neq \underline{v}$, that

$$T(\underline{u} - \underline{v}) \neq \underline{0}$$

This can be rephrased as: if $\underline{u} \neq \underline{v}$, which means $\underline{u} - \underline{v} \neq \underline{0}$, then $T(\underline{u} - \underline{v}) \neq \underline{0}$, and this in turn means that the nullspace of T is only $\underline{0}$. In fact, a vector $\underline{w} \triangleq \underline{u} - \underline{v}$ belongs to the vector space, so, $T\underline{w}$ must be different from $\underline{0}$!

So, the nullspace of a linear operator allows to characterize whether they are 1 to 1!

The other important subspace to be study to characterize an operator is the range of an operator.

$$\text{range}(T) = T(V) = \left\{ T\underline{v} \ : \ \underline{v} \in V \right\},$$

this is the set of all possible points which can be mapped can be reached by our operator! This, too, can be shown to be a subspace. In fact, if we multiply the input $\underline{v}$ times a scalar, a, then the output $T\underline{v}$ is just multiplied by a, and it still is in the range. Same for addition $\underline{y} + \underline{v}$!

The range determines if T is surjective. More in detail, if and only if range $(T) = V$, then T is surjective.
Since being surjective means just that the output is the entire space, this definition is also a "proof"
When V is finite dimensional, the dimension of V is equal to the dimension of the nullspace + the dimension of the range. Intuitively, if the nullspace has dimension K, which means, given n degrees of freedom in our input, T kills K of them, map them to 0, the remaining n-K degrees of freedom "affect the output". These are the n-K dimensions of the range.
The range is the orthogonal complement of the nullspace on V.
This theorem actually works not just for operators, but with general linear maps. Slightly more complicated but possible.
Considering for example

$$T = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

, the nullspace is $\operatorname{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right\}$,

the range is $\operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

This example makes seem this too simple. A much more horrible example is

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

, where $\operatorname{null}(T) = \operatorname{span}\left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$ ← they are equal!

$\operatorname{range}(T) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$ ←

Why?! Well, range is about the OUTPUT of the operator, nullspace is about the INPUT of the operator. They are different things!!!
There is no paradox here!

Some more detail about properties of operators: INVERTIBILITY!
A linear operator T has a left-inverse S if $ST = \mathbb{1}$.
" " " " " right " S' if $TS' = \mathbb{1}$

If it has both inverses, then they are the same:

$$S = S(\mathbb{1}) = S(TS') = (ST)S' = S'$$

In this case, T is said to be "invertible", and we define
$T^{-1} = S$

When do left and right inverses exist?
It turns out that the left inverse exists when, after we've applied T, we haven't lost information: an operator can restore the previous form.
Since destroying means being 1 on 1 (meaning, after I apply T I can get distinguishable values, so that if I have this value is in 1 on 1 correspondence with that I was coming from), the existence of left inverse means injectivity.

$\exists$ left inverse $\iff$ T is injective

For right inverse, the output of T had better to cover all the possibilities if we want to achieve identity by multiplying T by something on the right. Otherwise we could have zeros. So, surjectivity ;-)

$\exists$ right inverse $\iff$ T is surjective

If V is finite dimensional, T is injective if and only if it is surjective (think about null/range spaces), if and only if it is invertible.
This idea makes use of the fact that V is finite dimensional, and breaks down if this hypothesis fails. =C
For example, with ∞ dimensions, the sequence example: the left-shift operator is not injective:

$$L(x_1, x_2, x_3 \dots) = (x_2, x_3, x_4 \dots)$$

I can change x_1 as much as I want, and the result will not change.

On the other hand, R is injective and not surjective: the 1st component is always 0!

Not injective: Two inputs give the same output, meaning that we cannot go back from the output: lost information about "how to go back"

Not surjective: not all the outputs can be reached: lost information for this reason!

In ∞ dimensions we cannot look at range or nullspace: T is invertible if and only if it's injective and surjective!

Now, let us try to relate all of this to matrices!

Matrix form of an operator

An operator can be seen in a matrix form. However, to this aim, it is necessary to specify a basis. To define an operator and all concepts related to it, we need just a vector space and a mapping from it to itself.
To define a matrix, we need additional structure.
If we choose a basis $\{\underline{v}_1 \dots \underline{v}_N\}$, we can get a simple form of the operator, to work with. In fact, the fact that this is a basis means that any $\underline{v}$ can be written as

$$\underline{v} = a_1 \underline{v}_1 + a_2 \underline{v}_2 + \dots = \sum_{i=1}^{N} a_i \underline{v}_i ,$$

then, since T is linear,

$$T\underline{v} = T \sum_{i=1}^{N} a_i \underline{v}_i = \sum_{i=1}^{N} a_i \boxed{T\underline{v}_i}$$

Given one of these vectors, $\underline{v}_j$, it can be written as

$$\underline{v}_j = T_{1j}\underline{v}_1 + T_{2j}\underline{v}_2 + T_{3j}\underline{v}_3 + \dots T_{nj}\underline{v}_n = \sum_{i=1}^{n} T_{ij}\underline{v}_i \quad \left(\begin{array}{c}\text{expand one of}\\ \text{basis vectors}\end{array}\right)$$

So that by knowing T_{ij} I know the operator, how it acts on the basis.
In fact, given a generic $\underline{v}$,

$$T\underline{v} = \sum_{j=1}^{n} T a_j \underline{v}_j = \sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} a_j \underline{v}_i$$

So, if we know the expansion coefficients of our general vector $\underline{v}$ on the basis $\{\underline{v}_j\}$, and the T_{ij}, we can write this formula. This can be further manipulated as

$$\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} a_j \underline{v}_i = \sum_{i=1}^{n} b_i \underline{v}_i , \quad b_i = \sum_{j=1}^{n} T_{ij} a_j ,$$

which can be written in a more familiar form:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & \cdots \\ T_{21} & & \\ & & \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}$$

matrix form!

In particular, $\underline{\underline{T}}$ is the matrix form of the operator T!
Notice that we did not have axioms about matrix multiplication: we achieved this form from theory, just from the definition of vector space and of a basis!!

Similarly, we can derive matrix-matrix multiplication, considering two operators T and S, and we act on a vector of the basis, say $\underline{v}_k$:

$$TS\underline{v}_k = T\sum_j S_{jk}\underline{v}_j = \sum_i\sum_j T_{ij} S_{jk} \underline{v}_i$$

Now, the coefficients of $\underline{v}_i$ is

$$\sum_i \left[\sum_j T_{ij} S_{jk}\right] \underline{v}_i$$

this thing!
this is the matrix element of (TS),
this is the definition of matrix-matrix multiplication!

This is why we can think of matrices as a faithful representation of operators =)

Now.
We wanted to prove that all of these concepts can be defined within a very minimal framework. Yet, this is very important, since, even if matrices are very powerful tools, the fact that their employment requires the definition of a basis is not so nice. In fact, there is the risk to obtain results which rely on the properties of such basis: it would mean something like have a dependence on the coordinate system! o_o
For this reason, we are interested in quantities which are basis independent!
Basis independent quantities are: the trace, the determinant! Thus, because the trace turns out to equal the sum of the eigenvalues, and the determinant their product.
Basis-independent things are so, because they are functions of the eigenvalues! And don't care about their ordering.
PROBLEM SET these proofs!

Eigenvalues

They are basis-independent quantities. Eigenvalues come from a quite general idea, which is that of an "invariant subspace".
U is a T-invariant subspace if $T(U)$ (operator T, subspace U), means

$$T(U) = \{T\underline{u} \,,\, \underline{u} \in U\},$$ is contained in U!

Trivial
Examples: $\{\underline{0}\}$ subspace, V subspace (the entire space)

Focusing on 1-dimensional subspaces, we want subspaces composed by a SINGLE vector,

$$U = \text{span}\{\underline{u}\}$$

So, U being T-invariant is equivalent to

$$T\underline{u} \in U \iff T\underline{u} = \lambda\underline{u}$$

$T\underline{u}$ is some MULTIPLE of $\underline{u}$! This is the eigenvector equation!
λ is the eigenvalue, $\underline{u}$ is the eigenvector
Trivial, $\underline{u} = \underline{0}$, not interesting.
So, λ is an eigenvalue of T if $T\underline{u} = \lambda\underline{u}$, for $\underline{u} \neq \underline{0}$.
The spectrum of T is the collection of all eigenvalues!

Examples

Given $V = \mathbb{R}^3$, T = rotation around $\hat{z}$ axis (θ-rotation),

$$T = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

An eigenvalue of this matrix is 1, with $\underline{v} = \begin{bmatrix} 0 \\ 0 \\ \vdots \end{bmatrix}$.

Then, it has two more complex eigenvalues.
Another example:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

now, T has NO eigenvalues.

William Detmold

Lecture 7: Linear Algebra: Vector Spaces & Operators (continued)

Now we will complete the discussion about eigenvalues/eigenvectors
We were talking about T-invariant subspaces. A subspace is T-invariant if

$$T(U) = \{T\underline{u} : \underline{u} \in U\} \subseteq U$$

But we said that we focus on 1D subspaces. So, we can write U as

$$U = \{\lambda\underline{u} : \lambda \in \mathbb{F}\}$$

In this case we have a very simple equation, which is

$$T\underline{u} = \lambda\underline{u} \qquad \text{T-invariant, 1-D.} \qquad (1)$$

This is just an "abstract" version of what we've discussed before, with matrices and vectors.
The spectrum of an operator is the set of all its eigenvalues.

So, if λ is an eigenvalue, (1) tells us that

$$(T - \lambda\mathbb{1})\underline{u} = \underline{0} \qquad (2) \qquad \text{(just taking the RHS to the left)}$$

So, this is an operator now, as it is "multiplying" $\underline{u}$. This equation (2) tells us that the operator $(T-\lambda\mathbb{1})$ is not injective, as it is mapping the vector $\underline{u}$ to $\underline{0}$, "if applied to" $\underline{u}$. Then, we can write its nullspace!

$$\text{null}(T-\lambda\mathbb{1}) = \{\text{set of eigenvectors of } T \text{ with eigenvalue } \lambda\}$$

So, every eigenvector (of T) with eigenvalue λ causes $(T-\lambda\mathbb{1})\underline{u}$ to go to $\underline{0}$. As discussed in the previous lecture, if this operator is not injective, then it is not invertible. This allows us to write the following theorem.

Let $T \in \mathcal{L}(V)$, and $\{\lambda_i\}$ its (distinct) eigenvalues, with the corresponding eigenvectors $\{\underline{u}_i\}$, $i = 1 \dots n$. Then, the list $\{\underline{u}_1 \dots \underline{u}_n\}$ is a linearly independent set

(proof)

By contradiction. If it was false, then there is a nontrivial relation $c_1\underline{u}_1 + c_2\underline{u}_2 + \dots + c_K\underline{u}_K = \underline{0}$, where the $\{c_i\}$ cannot be all equal to $\underline{0}$.
Let there be at least $K \leq n$ such that this occurs. Then, let us apply our operator $(T-\lambda_K\mathbb{1})$:

$$(T-\lambda_K\mathbb{1})(c_1\underline{u}_1 + c_2\underline{u}_2 + \dots + c_K\underline{u}_K) = c_1(\lambda_1-\lambda_K)\underline{u}_1 + \dots + c_{K-1}(\lambda_{K-1}-\lambda_K)\underline{u}_{K-1} + \underline{0}\,\underline{u}_K = \underline{0}$$

(the step is: $(T-\lambda_K\mathbb{1})c_1\underline{u}_1 = Tc_1\underline{u}_1 - \lambda_K c_1\underline{u}_1 = \lambda_1 c_1\underline{u}_1 - \lambda_K c_1\underline{u}_1$)

So. This equation involves less than K terms, and then, the space of linearly

independent terms is smaller than that of the hypothesis, so we reached a contradiction. There must be no nontrivial relation.

Re-explain

Objective: verify that eigenvectors are linearly independent.

How: by contradiction: assume that they are NOT independent and see if some contradiction arises. [where K is the minimum dimension such that I can achieve linear dependences]

Let us fix $K \leq n$. Then, if the eigenvectors are NOT linearly independent, then there must be a combination nontrivial (not all coefficients 0),

$$c_1 \underline{v}_1 + c_2 \underline{v}_2 + \dots + c_K \underline{v}_K = \underline{0} \qquad (3)$$

We focus just on K eigenvectors. Then, by applying $(T - \lambda_K \mathbb{I})$, we obtain

$$\underbrace{c_1(\lambda_1 - \lambda_K)}_{d_1} \underline{v}_1 + \underbrace{c_2(\lambda_2 - \lambda_K)}_{d_2} \underline{v}_2 + \dots + \underbrace{c_{K-1}(\lambda_{K-1} - \lambda_K)}_{d_{K-1}} \underline{v}_{K-1} + \underbrace{0}_{d_K} \underline{v}_K$$

this shows that, if K was the minimum dimension of the space of linearly dependent vectors (hypothesis), by applying this procedure, I obtain

$$d_1 \underline{v}_1 + d_2 \underline{v}_2 + \dots + d_{K-1} \underline{v}_{K-1} = \underline{0},$$

where all the $\{d_K\}$ are surely $\neq 0$ for the hypothesis $\lambda_i \neq \lambda_j$, $c_i \neq 0$. So, this implies also that K-1 vectors are exhibiting dependences and this is the contradiction! Contradiction with the "least", "minimum" hypothesis.

[SCRIVERE UNA PROOF MIGLIORE]

Another theorem

Let $T \in \mathcal{L}(V)$, with V being a finite-dimensional complex vector space, then there is at least 1 eigenvalue!

Notice that the last example of previous lecture had NO eigenvalues. But, we were focusing on real spaces!

Everything we stated until now for eigenvalues and eigenvectors was not referred to a particular basis. In fact, they are basis independent. But, in view of obtaining numbers we could choose a basis and then introduce the matrix representation of the operator. Then, since for an eigenvalue λ the operator $(T - \lambda \mathbb{I})$ is not invertible, then NO ONE OF ITS MATRIX REPRESENTATIONS will be invertible. In any basis!

If a matrix is not invertible, then its determinant is 0. So, since we can write the determinant of a matrix representation of $(\underline{\underline{T}} - \lambda \underline{\underline{\mathbb{I}}})$, we find that it is a polynomial function of λ:

$$f(\lambda) = \det\left(\underline{\underline{T}} - \lambda \underline{\underline{\mathbb{I}}}\right) = 0 \qquad \text{characteristic polynomial!}$$

$$= (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n) \qquad \text{(in factor form)},$$

where the zeros $\{\lambda_i\}$ are in general complex and can be repeated. The extreme situation is: all eigenvalues equal, degenerate, to λ_1: $(\lambda - \lambda_1)^n$. DEGENERACY.

Now, let us change subject

Inner products

First. What is an inner product? It is a very specific map. Given a vector space V, it is a map from $V \otimes V \to \mathbb{F}$; it is also called "dot product": it is its generalization

For the dot product, let $V = \mathbb{R}^n$, so $\underline{a} = (a_1 \dots a_n)$

$\underline{b} = (b_1 \dots b_n)$,

then we can define the dot product $\underline{a} \cdot \underline{b}$ as

$$\underline{a} \cdot \underline{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + b_2 a_2 + \dots + a_n b_n \in \mathbb{R}$$

This is an example of inner product =)

Which are the properties of the inner product?

(1) $\underline{a} \cdot \underline{b} = \underline{b} \cdot \underline{a}$ (commutative)

$\underline{a} \cdot \underline{a} \geq 0$ (it is interpreted as a norm: length: greater than 0, and

(2) $\underline{a} \cdot \underline{a} = 0 \iff \underline{a} = \underline{0}$

(3) $\underline{a} \cdot (\beta_1 \underline{b}_1 + \beta_2 \underline{b}_2) = \beta_1 \underline{a} \cdot \underline{b}_1 + \beta_2 \underline{a} \cdot \underline{b}_2$, $\beta_1, \beta_2 \in \mathbb{R}$

(4) As mentioned in (2), the dot product defines a length, or "norm", of a vector:

$$\|\underline{a}\|^2 \triangleq \underline{a} \cdot \underline{a}$$

(5) Schwarz inequality: $|\underline{a} \cdot \underline{b}| \leq \|\underline{a}\| \, \|\underline{b}\|$

Now, this was true for $\mathbb{R}^n$. What about our true focus, which is, a complex vector space?

In a complex vector space, the first problem is defining what we intend by a "norm": we have no guarantee that $\underline{v}\cdot\underline{v}$ is real!
Then. We want a good definition for a length.
Let $\underline{z} \in \mathbb{C}^n = (z_1, z_2, z_3, z_4, \dots z_n)$, $z_i \in \mathbb{C}$.
Then, for a single complex number, it is known that its "length" is its absolute value: the distance from the origin of the complex plane. Since $\mathbb{C}^n$ is more complicate, let us think of $\|\underline{z}\|$ as

$$\|\underline{z}\| = \sqrt{|z_1|^2 + |z_2|^2 + \dots + |z_n|^2} = \sqrt{z_1^* z_1 + z_2^* z_2 + \dots + z_n^* z_n}$$

So, we could consider introducing, in this "generalized dot product" (the inner product), complex conjugation. We want that our inner product produces a length!
So, this generalized inner product should satisfy the following axioms: we want to maintain the dot product properties, and it should be a map from $V \otimes V$ to $\mathbb{F}$

$$\langle \cdot | \cdot \rangle : V \otimes V \longrightarrow \mathbb{C},$$

linear

$$\langle \underline{v} | \underline{b} \rangle \quad \underline{v}, \underline{b} \in V, \quad \langle \underline{v} | \underline{b} \rangle \in \mathbb{C}$$

Axioms

(1) $\langle \underline{v} | \underline{b} \rangle = \langle \underline{b} | \underline{v} \rangle^*$

(2) $\langle \underline{v} | \underline{v} \rangle \in \mathbb{R}$, and ≥ 0, $= 0$ only for $\underline{v} = 0$

(3) we want distributivity:

$$\langle \underline{v} | \beta_1 \underline{b}_1 + \beta_2 \underline{b}_2 \rangle = \beta_1 \langle \underline{v} | \underline{b}_1 \rangle + \beta_2 \langle \underline{v} | \underline{b}_2 \rangle \quad \beta_1, \beta_2 \in \mathbb{C}$$

How can we define the inner product satisfying these axioms?
Notice that the notation $\langle \cdot | \cdot \rangle$ has been proposed by Dirac (Dirac notation), in math literature $(\cdot, \cdot)$ or $\langle \cdot, \cdot \rangle$ is used. They mean the same thing.
Another axiom we could demand is

$$\langle \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 | \underline{b} \rangle = \alpha_1^* \langle \underline{v}_1 | \underline{b} \rangle + \alpha_2^* \langle \underline{v}_2 | \underline{b} \rangle$$

So, there is an asymmetry! This, still, follows from (1)! ((1)+(3))
The complex conjugates on α are very important. Let us prove this

$$\langle \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 | \underline{b} \rangle = \langle \underline{b} | \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 \rangle^* = \left(\alpha_1 \langle \underline{b} | \underline{v}_1 \rangle + \alpha_2 \langle \underline{b} | \underline{v}_2 \rangle\right)^* =$$

$$= \alpha_1^* \langle \underline{b} | \underline{v}_1 \rangle^* + \alpha_2^* \langle \underline{b} | \underline{v}_2 \rangle^* = \alpha_1^* \langle \underline{v}_1 | \underline{b} \rangle + \alpha_2^* \langle \underline{v}_2 | \underline{b} \rangle$$

Example (1) : finite-dimensional

$V = \mathbb{C}^n$, so

$$\langle \underline{v} | \underline{b} \rangle = v_1^* b_1 + v_2^* b_2 + \dots = \sum_{i=1}^{n} v_i^* b_i, \qquad (4)$$

so the only difference from dot product in real vector spaces is complex conjugation.

Example (2) : ∞-dimensional

Let V be the set of all $f(z) \in \mathbb{C}$ with $z \in [0, L]$
A natural norm for this space is: (DEFINITION)

$$f, g \in V, \qquad \langle f | g \rangle = \int_0^L f^*(x) g(x) dx$$

this integral can be seen as the limit of the sum (4), it's generalization!

So, previously, until this moment, we had just a vector space. Now, we have a vector space with an inner product. This is something stronger, and it allows us to do things that we weren't able to do before!!!
For instance, now we can finally talk about orthogonality!
In fact, we weren't basically inable to ask questions about 2 objects in the vector space!
Given $\underline{v}, \underline{b} \in V$ two elements of our complex vector space V, given their inner product $\langle \underline{v} | \underline{b} \rangle$, they are orthogonal if
$\langle \underline{v} | \underline{b} \rangle = 0$!
this is the generalization, for an abstract vector space, of the $\mathbb{R}^n$ concept of orthogonality!
If we have a set of vectors $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ such that $\langle \underline{e}_i | \underline{e}_j \rangle = \delta_{ij}$, this set is said to be "orthonormal"
Then, we can define, compute, the components of vectors, in a basis-dependent way. Let $\{\underline{e}_i\}$ be a basis in V.

If we demand that they are orthonormal, then we can write a vector $\underline{v} \in V$ as

$$\underline{v} \in V \Rightarrow \underline{v} = \sum_{i=1}^{n} v_i \underline{e}_i \quad \text{this even if } \{\underline{e}_i\} \text{ are not orthonormal.}$$

BUT, if they are orthonormal, we have some advantages. In fact,

$$\langle \underline{e}_k | \underline{v} \rangle = \langle \underline{e}_k | \sum_{i=1}^{n} v_i \underline{e}_i \rangle = \sum_{i=1}^{n} v_i \langle \underline{e}_k | \underline{e}_i \rangle = \sum_{i=1}^{n} v_i \delta_{ki} = v_k$$

This allows to compute the component of $\underline{v}$! This, thanks to the inner product! This is the added value of having an inner product!

We can also talk about norm:

$$\|\underline{v}\|^2 = \langle \underline{v} | \underline{v} \rangle,$$

which is a sensible norm, thanks to axiom 2. This allows to generalize also Schwarz's inequality,

$$|\langle \underline{v} | \underline{b} \rangle| \leq \|\underline{v}\| \, \|\underline{b}\|$$

we can also write down a triangle inequality,

$$\|\underline{v} + \underline{b}\| \leq \|\underline{v}\| + \|\underline{b}\|$$

We are doing a lot of math. What is it aimed to?!
We can try to discuss, within this framework, the space where our quantum states live in. Basically, they are vector spaces having an inner product, so "inner product spaces", or "Hilbert spaces".
This is quite straightforward with finite dimensions, not so much with ∞!
For ∞ dimensional vector spaces with an inner product, we have to make sure that such space is complete. A complete vector space with inner product is again a Hilbert space.
What does it means "complete"? See notes of Functional Analysis.
Polynomial example, min. 53/54

How can we obtain an orthonormal basis? From the Gram-Schmidt procedure!
Quite similar to $\mathbb{R}^3$, just with the new inner product.
Let $\{\underline{v}_1, \underline{v}_2 \dots \underline{v}_n\}$ linearly independent vectors, then we can construct another list $\{\underline{e}_1, \underline{e}_2 \dots, \underline{e}_n\}$, which is orthonormal, as:

$$\underline{e}_j = \frac{\underline{v}_j - \sum_{i<j} \langle \underline{e}_i | \underline{v}_j \rangle \underline{e}_i}{\left\| \underline{v}_j - \sum_{i<j} \langle \underline{e}_i | \underline{v}_j \rangle \underline{e}_i \right\|}$$

Finally. Inner products can be used to find the orthogonal complement of whatever space we want.
Let V be a vector space, and U a set of $\underline{v} \in V$, for instance, given $\mathbb{R}^3 \equiv V$,
$U \equiv \mathbb{R}^2$.
Then, we can define the orthogonal complement $U^\perp$ as

$$U^\perp = \{ \underline{v} \in V : \langle \underline{v} | \underline{u} \rangle = 0, \forall \underline{u} \in U \}$$

everything in U is orthogonal to $U^\perp$.
$U^\perp$ is a subspace (vector) of V.

Then, Theorem
If U is a subspace, then $V = U \oplus U^\perp$: $\oplus$ is the direct sum.
Thus familiar

Now, we will say more about...

Dirac notation

We started discussing these objects
$|b\rangle$
and calling them "kets". These things live in our vector space V:
they are basically vectors
If we look at how we write the inner product, well, this looks a lot like the right half!
So,
$\langle a | b \rangle$: it's like breaking the inner product into 2 pieces.
$\langle a |$ bra, $| b \rangle$ ket
What Dirac said was: "well, if $\langle \quad \rangle$ is a bracket, then by breaking like $\langle | \quad | \rangle$ we'll
Kets live in V, in the vector space.

Instead the bras are not vectors in V. So, $\langle a|$ is something different! The bra takes a vector, $|b\rangle$, and returns a number!
We have to think of the inner product as the result of the action of the bra on the ket! So, $\langle a|$ is actually a map! A map from V to $\mathbb{C}$!
So, $\langle a| \notin V$!!! Instead, $\langle a|$ belongs to what is called "dual space" V^*. $\langle a|$ is a linear map.
What is V^*? V^* is the space of linear maps from V to $\mathbb{C}$, but itself it is a vector space, so that addition of those maps is defined, as well as scalar multiplication!
So, we can define a bra $\langle w|$ such as:

$$\langle w| = \alpha\langle a| + \beta\langle b|, \qquad \langle a|, \langle b|, \langle w| \in V^*$$

We define it such that

$$\langle w|v\rangle = \alpha\langle a|v\rangle + \beta\langle b|v\rangle, \quad \forall w \in V,$$

So, for all vectors $v \in V$, w is defined such that this relation holds.
This tells us that V^* is a vector space.

There is a correspondence between the vectors living in the space V and the elements of the dual space V^*.
So, for any $|v\rangle \in V$, there is a unique bra $\langle v| \in V^*$;
such uniqueness can be shown by contradiction. In fact, assume that there exist $\langle v|, \langle v'|$ such that $\langle v|u\rangle = \langle v'|u\rangle, \forall |u\rangle \in V$.
Then, we can construct

$$\langle v|u\rangle - \langle v'|u\rangle = 0 = (\langle u|v\rangle - \langle u|v'\rangle)^* = (\langle u|v-v'\rangle)^* =$$

$$= \langle v-v'|u\rangle \overset{?}{=} 0,$$

yes, but since this must be true $\forall |u\rangle \in V$, then $\langle v-v'| = 0$ is the only way! $\boxed{v=v'}$ → contradiction

This proves that we have a 1 to 1 correspondence between elements in dual and vector spaces!
More generally, there is a correspondence between

$$|v\rangle = a_1|a_1\rangle + a_2|a_2\rangle \quad \text{and} \quad \langle v| = a_1^*\langle a_1| + a_2^*\langle a_2|$$

Any linear map $V \to \mathbb{C}$ defines a bra! So, there is an element of V^* corresponding to it.
More concretely, we can think of kets as column vectors, bras are sort of "row" vectors (conjugate), and think of inner product just as a matrix multiplication!

$$\langle a| \leftrightarrow \begin{bmatrix} a_1^* & a_2^* & a_3^* & \cdots & a_n^* \end{bmatrix}, \qquad |v\rangle \longleftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix}$$

$$\langle a|v\rangle = \sum_{i=1}^{n} a_i^* v_i$$