Introduction to Quantum Theory

By R.C. Dennis

Introduction

We shall begin our study of Quantum Theory (QT) from a completely abstract point of view. We shall build up the abstract machinery of QT and slowly make connections with the physical world around us. The is a very platonic approach whereby we begin with perfect and ideal structures in the realm of ideals and then grope away at the natural world via the various mathematical tendrils we shall develop. This has both benefits and shortcomings and i hope that you find the benefits outweigh those shortcomings. Typically courses take the opposite approach and begin the study of the quantum realm via experiments such as the two-slit interference pattern, the ultra-violet catastrophe, Mach-Zehnder interferometry, or the discrete spectra of atomic energy transitions. Indeed we shall discuss these realities in turn but only after taking a high-level and abstract approach.

The strategy is:

1. Define what a "quantum object" or "state" is. By "object" i dont necessarily mean physical matter. I use "object" here in an abstract and mathematical sense.

2. Postulate that a specific mathematical operation on these "states" is related to the probability of some experimental outcome or observation. The is called the Born rule and has yet to be proven definitively, so at best we will hold it as an axion of our inquiries.

3. Define an "operator" which is a map from one state to another. The previous abstract considerations will produce a rigid structure restricting the types of operators can apply in QT.

4. Define the various symmetries we expect of our universe. For example, outcomes shouldn't depend on where you are in the universe all other things held constant. This would be spatial symmetry and we will build an operator from it that translates states in space.

5. If we demand certain symmetries to hold then we will find rigorously special states and operators on those states to determine outcomes of experiments.

6. We shall define "observables" and "measurements".

7. Finally we shall find the Schrödinger equation which determines the evolution of these quantum states. The we will apply these ideas to our physical world.

Quantum States:

We shall define a quantum state to be a type of abstract vector inside a specific abstract space. This is a mathematical "vector space" and has nothing, yet, to do with our familiar physical space. viz, (x, y, z). Often words will overlap between mathematics and physics and we should be mindful of those words. Another word is "field" which means different things to a mathematician and to a physicist. Context matters and here we are within the platonic realm of ideas.

We shall write the vectors using the following notation. General vectors will have greek letters whereas specific vectors will often have latin letters.

$$|\psi\rangle, |\phi\rangle, |\xi\rangle, \cdots$$

and they live in a place called "Hilbert Space". we write:

$$|\psi\rangle \in \mathcal{H}$$

At this point it's merely notational. For now nevermind what "Hilbert" means: at this point it's simply a "vector space" with some dude's name. It will have special properties we shall detail a little later.

Just like $\vec{x}$ is a notation....they're both vectors. However, unlike our classical physics vectors $\vec{x}$ which hold $\mathbb{R}$ values, our quantum vectors $|\psi\rangle$ will be $\mathbb{C}$ valued. In quantum theory its traditional to use Greek letters for general or arbitrary states, and Latin letters for specific states. We will almost always be using ψ as the cultural tradition is to do so.

Okay, if these quantum vectors are complex valued then we can "represent" them as columns of complex numbers with respect to a chosen basis. We're used to doing this without much thought. For example, classically we decompose a vector in terms of our coordinate system of our choosing. viz,

$$\vec{r} = r_x\hat{x} + r_y\hat{y} + r_z\hat{z} = \sum r_i\hat{x}^i$$

where

$$\hat{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \hat{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \hat{z} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

therefor we write $\vec{r}$ as

$$\vec{r} = \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix}$$

This is what we've always done without all the gory details! \

Similarly, we will write a state vector $|\psi\rangle$ in terms of a basis. Importantly, this basis has NOTHING to do with (x, y, z) or space or anything like that. It's entirely an abstract basis.

We say, "if we know $|\psi\rangle$ is a vector in some vector space $\mathcal{H}$ then we can choose a basis of $\mathcal{H}$ such that $|\psi\rangle$ is a linear combination of those basis vectors".

That's exactly what we do with choosing coordinate systems in classical physics. Okay, let's pick a random letter to use as our basis vectors of $\mathcal{H}$: say, $|n\rangle$. Then,

$$|\psi\rangle = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \end{bmatrix}$$

and

$$|\psi\rangle = \sum_n \alpha_n |n\rangle$$

where $\alpha \in \mathbb{C}$ are the complex components of our vector (just like the r_x, r_y, r_z of $\vec{r}$). So we have vectors then what about matrices? Hold your

horses, we've defined column vectors but for every vector space there exists a "dual" vector space.

Your whole life you've been using two vector spaces but weren't aware of it. Dear reader, i ask you, "how do you take a dot product?"

you respond, "easily, take two vectors and multiply their components and add them all up". Yes, but how do you do that with two column vectors? You never realized it but doing the dot product amounts to taking a row vector and multiplying by a column vector using the typical rules of matrix multiplication. viz, 1×3 matrix multiplied by a 3×1 matrix produces a 1×1 martrix (otherwise known as a number!). So to do the dot product i have to take a column vector and tip it over into a row vector.

$$\vec{a} \cdot \vec{b} = \sum_i a_i b_i$$

often when we have repeated indices we dont write the $\sum$ symbol and agree to sum over indices when they appear twice. In these notes ill write out the full sums but you should try and get used to the notation without the sum symbol because its VERY VERY VERY useful. viz,

$$\vec{a} \cdot \vec{b} = a_i b_i$$

okay, as matrices we write this operation on column vectors as

$$\vec{a}^T \vec{b} = \vec{a} \cdot \vec{b}$$

Where T means "transpose" otherwise known as switching rows and columns.

$$\vec{a} \cdot \vec{b} = [a_1 a_2 \ldots.] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix} = a_1 b_1 + a_2 a_2 + \ldots = \sum_{i=0}^{N} a_i b_i$$

This is how we do dot products in classical physics land. Our vector space is composed of column vectors and our "dual" vector space is composed of "row" vectors. For every column there is a row and vice versa and the T - transpose maps from our vector space to our "dual vector space. The row vectors define a new mathematical object called a "Linear Functional". A Linear Functional is a map that eats a vector and spits out a (real or complex) number. Yes, yes, i know this is overkill for classical physics, but, dear reader, it is EXTREMELY important in quantum theory but especially in differential geometry and gravity.

How does this work for our quantum Hilbert Space? Everything will work out the same, except we have complex numbers. It turns out the proper thing to do is to take the complex conjugate of our vector's elements and then also transpose. This called the Hermitian conjugate. We define our "inner product" as;

for $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$

$$\langle\alpha|\beta\rangle = [\alpha_1^*\alpha_2^*, \ldots.] \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \end{bmatrix} = \alpha_1\beta_1 + \alpha_2^*\beta_2 + \ldots = \sum_{i=0}^{N} \alpha_i^*\beta_i$$

where $\langle\alpha| = (|\alpha\rangle)^\dagger$ where $\dagger$ (said as "dagger") means to take the complex conjugate AND transpose.

$

$$\langle\alpha| = (|\alpha\rangle)^\dagger = [\alpha_1^*\alpha_2^*, \ldots.]$$

Now we're cooking. We have imposed a structure onto our Hilbert Space called an "inner product". It wasnt there before, we created it and said "this is something we'd like to do".

Now we can discuss things like "what is the length of a quantum vector?"

$$\langle\alpha|\alpha\rangle = [\alpha_1^*\alpha_2^*, \ldots.] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \end{bmatrix} = \alpha_1^*\alpha_1 + \alpha_2^*\alpha_2 + \ldots$$

$$= |\alpha_1|^2 + |\alpha_2|^2 + \ldots$$

So we have notions of length. We know not what it means physically but we've created it.

NOTATION ALERT

Mathematicians like to write the dot/inner product as $(\alpha, \beta) = \langle\alpha|\beta\rangle$. I personally don't like it and think the "bracket" notion is much clearer but what can you do? You'll come across both notations.

Linear Operators

An operator T is yet another specific (and new to us) mathematical object that maps vectors into vectors, viz

$$\hat{T} : \mathcal{H} \to \mathcal{H}$$

as

$$\hat{T}|\alpha\rangle = |\beta\rangle$$

You've seen such things: if we choose a basis of $\mathcal{H}$ then it's simply a matrix. Explicitly,

$$|\beta\rangle = \hat{T}|\alpha\rangle \Rightarrow \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & \cdots \\ T_{21} & T_{22} & \cdots \\ \vdots & & \\ T_{n1} & \cdots & T_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \vdots \end{bmatrix}$$

In quantum theory we generally place hats on "operators". An operator is something we can do to a quantum state such that we get a different state. This is beginning to sound physical

Operators are LINEAR, i.e.

$$\hat{T}(c_1|\alpha_1\rangle + c_2|\alpha_2\rangle) = \hat{T}c_1|\alpha_1\rangle + \hat{T}c_2|\alpha_2\rangle = c_1(\hat{T}|\alpha_1\rangle) + c_2(\hat{T}|\alpha_2\rangle)$$

$$\therefore \hat{T}(c_1|\alpha_1\rangle + c_2|\alpha_2\rangle) = c_1|\beta_1\rangle + c_2|\beta\rangle$$

This is what we mean by "Linear Operator".

Theorem: If the dimension of Hilbert Space is finite then we can represent operators as matrices.

Consider

$$\hat{T}|\alpha\rangle = |\beta\rangle$$

First, let's pick a complete and orthonormal basis $|n\rangle = |0\rangle, |1\rangle, |55\rangle$,etc. Orthonormality means that our basis vectors are of unit length and all perpendicular to each other (just like our cartesian coordinates). Orthonormal bases are super-handy and we try to use them where ever we can!
viz,

$$|\alpha\rangle = \sum_j \alpha_j |j\rangle$$

$$|\beta\rangle = \sum_k \beta_k |k\rangle$$

Dear reader, it doesnt matter what we call an index when we sum over it and i intentionall chose two different indices j and k. they both range from 0 to N where N is the dimension of the space $\mathcal{H}$. so,

$$\hat{T}\sum_j \alpha_j|j\rangle = \sum_k \beta_k|k\rangle =$$

Now we will come from the side and hit the expression with $\langle i|$

$$\langle i|\hat{T}\sum_j \alpha_j|j\rangle = \langle i|\sum_k \beta_k|k\rangle$$

Now what? well notice that since $\hat{T}$ is a linear operator it distributes into the sum and bypasses the α_j and β_k's. That is, we get

$$\sum_j \alpha_j\langle i|\hat{T}|j\rangle = \sum_k \beta_k\langle i|k\rangle$$

but, by orthonormality

$$\langle i|k\rangle = \delta_{ik}$$

$\delta_{ik} = 0$ if $i \neq k$ and $\delta_{ik} = 1$ if $i = k$

Therefore the sum over k is entirely zero except where $i = k$ where the sum evaluates to β_i

$$\sum_j \alpha_j\langle i|\hat{T}|j\rangle = \beta_i = \sum T_{ij}\alpha_j = T_{ij}\alpha_j$$

in the last expression sum over j is implied apropos our summation convention of not writing $\sum$ everywhere.

$\langle i|\hat{T}|j\rangle = T_{ij}$ are called "The Matrix Elements" of the operator $\hat{T}$ in basis $|n\rangle$. for each value of i and j T_{ij} is simply some complex number. $\hat{T}$ itself, in this basis, is a matrix of complex numbers.

As you get comfortable with Einstein's summation convention you'll come to view the expression $\beta_i = T_{ij}\alpha_j$ as a vector equals a matrix times a vector without much though. Matrix elements will be VERY important for us.

Now, if we can do this special "dagger" operation to turn a column vector into a row vector then how much an operator respond under the hermitian adjoint?

$$\hat{T}|\alpha\rangle = |\beta\rangle \Rightarrow (|\beta\rangle)^\dagger = \langle\beta|$$

then

$$(\hat{T}|\alpha\rangle)^\dagger = \langle\beta| \Rightarrow (\hat{T}|\alpha\rangle)^\dagger = \langle\alpha|T^\dagger$$

our adjoint (dagger) switches the order of a product just like the transpose does for real matrices. viz

$$(ABCD)^\dagger = D^\dagger C^\dagger B^\dagger A^\dagger$$

additionally dagger acts as

$$(cA)^\dagger = c^* A^\dagger$$

$$(A+B)^\dagger = A^\dagger + B^\dagger$$

furthermore, you may prove that

$$(\langle\psi|\hat{T}^\dagger|\phi\rangle)^* = \langle\phi|\hat{T}|\psi\rangle$$

and

$$\langle\psi|\phi\rangle^* = \langle\phi|\psi\rangle$$

and

$$\langle\hat{T}^\dagger\psi|\phi\rangle = \langle\phi|\hat{T}\psi\rangle$$

where $\langle\hat{T}^\dagger\psi| = (\hat{T}|\psi\rangle)^\dagger$ and $|\hat{T}\psi\rangle = \hat{T}|\psi\rangle$ are the transformed vectors.

Since these operators are matrices they have determinants and traces.

$$\det T = \prod_i t_i$$

$$\mathrm{Tr}\, T = \sum_i T_{ii} = \sum_i t_i$$

where t_i are the eigenvalues of $\hat{T}$ which we may obtain in the standard way. where we solve,

$$\det(T - t\mathbb{1}) = 0$$

for t. In general there will be at most $\dim \mathcal{H}$ eigenvalues.

Self-Adjoint or Hermitian Operators

You may have already asked yourself the question: "does there exist operators $\hat{A}$ such that $\hat{A} = \hat{A}^\dagger$?"

This would be a special class of operators indeed. Operators that satisfy

$$\langle\phi|\hat{T}|\psi\rangle = \langle\psi|\hat{T}|\phi\rangle^*$$

are called "Self-Adjoint" and a matrix given by $A_{ij} = A^*_{ji}$ is called Hermitian. We will refer to operators as "Hermitian" rather than make such picky differences.

So, we have Hermitian Operators.

Theorem: Hermitian matrices have real valued eigenvalues
Proof:

$|t_1\rangle$, $|t_2\rangle$ are eigenvectors of $\hat{T}$ if

$$\hat{T}|t_i\rangle = t_i|t_i\rangle$$

then

$$\langle t_1|\hat{T}|t_1\rangle = \langle t_1|\hat{T}|t_1\rangle^* = 0$$

$$t\langle t_1|t_1\rangle = t_1^*\langle t_1|t_1\rangle = 0 \Rightarrow t = t^* \in \mathbb{R}$$

Theorem: Eigenvectors of Hermitian matrices are orthogonal

Proof:
$|t_1\rangle$, $|t_2\rangle$ are eigenvectors of $\hat{T}$ if

$$\hat{T}|t_i\rangle = t_i|t_i\rangle$$

where t_i are distinct eigenvalues. Then, from our definition of self adjoint

$$\langle t_1|\hat{T}|t_2\rangle - \langle t_2|\hat{T}|t_1\rangle^* = 0$$

$$t_2\langle t_1|t_2\rangle - t_1\langle t_2|t_1\rangle^* = 0$$

$$(t_2 - t_1)\langle t_2|t_1\rangle^* = 0$$

$$\therefore \langle t_2|t_1\rangle = 0$$

Completeness:
We shall say a set of basis vectors $|n\rangle$ are complete if

$$|\psi\rangle = \sum_n \alpha_n |n\rangle$$

and

$$\sum_i |n\rangle\langle n| = \mathbb{1}$$

i.e., this particular sum over states produces the identity matrix. This operation of a column vector times a row vector is called an "outer product".

$$|\beta\rangle\langle\alpha| = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \end{bmatrix} [\alpha_1^* \alpha_2^*, \ldots.]$$

produces a matrix.

Anytime you are trying to solve a quantum problem you may find it useful to multiply operators by $\mathbb{1}$ just like in algebra we sometimes multiply expressions by different representations of the number 1. This is a very handy trick we will leverage a LOT.

Next, if $\hat{T}|t_i\rangle = t_i|t_i\rangle$, then we can "spectrally decompose" $\hat{T}$ as

$$\hat{T} = \sum_i t_i |t_i\rangle\langle t_i|$$

This is a very nifty way to handle operators. It amounts to decomposing an operator/matrix in a basis where the matrix is diagonal. That's all this expression says. In our $|t\rangle$ eigenbasis we have

$$\hat{T} = \begin{bmatrix} t_1 & 0 & \cdots \\ 0 & t_2 & \cdots \\ \vdots & & \\ 0 & \cdots & t_n \end{bmatrix}$$

a tidy diagonal matrix. This is also called "diagonalizing an operator". Solving a quantum system usually amounts to diagonalizing an operator.

Another important tool is, given some function f we can find $f(\hat{T})$ as

$$f(\hat{T}) = \sum_i f(t_i) |t_i\rangle\langle t_i|$$

Handling infinite dimensional Hilbert Spaces

So far, all our work has been towards finite dimensional complex vector

spaces. The land of column vectors and matrices. However, we can, AND DO, have infinite dimensional vector spaces and the question becomes "how to handle these".

An example of an operator acting on an infinite dimensional vector space would be the derivative where functions would be the "vectors". Youve seen this idea without realizing it: The Taylor Expansion of a function. The polynomials $(1, x, x^2, \ldots)$ are the basis "vectors" and a general "vector" can be written as a linear combination of these bases vectors. Functions can behave perfectly well as vectors. Another example, which we will learn in detail later, is the Fourier Series with basis $(1, \sin x, \cos x, \ldots)$.

For example, in the vector space of polynomials we may expand any polynomial in terms of basis vector.

e.g. $f(x) = 3 + x^3 - 64x$ represented as a vector would be

$$\begin{bmatrix} 3 \\ -64 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

In the basis spanned by $(1, x, x^2, ...)$

The derivative of this function is $f'(x) = -64 + 3x^2$, given by the vector

$$\begin{bmatrix} -64 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

Your task: Represent the operator $\frac{d}{dx}$ as an infinite matrix.

Clearly, we have the ability to deal with infinite matrices all along! In QT we will expand what we normally think of as vectors to also include functions and differential operators. A general differential operator might be something like

$$\hat{D} = 8\frac{d}{dx} + 5e^{\frac{d}{dx}}$$

You naturally ask "how would i do the exponential of a derivative?" by Taylor expanding! Call $\frac{d}{dx} = \hat{p}$. Then

$$e^{\hat{p}} = 1 + \hat{p} = \frac{1}{2!}\hat{p}^2 + \cdots$$

Okay, what does $\hat{D}$ act on? "Vectors, of course":

$$\hat{D}|\psi\rangle = (8\frac{d}{dx} + 5e^{\frac{d}{dx}})|\psi\rangle$$

Now, this is all incredibly abstract, but it almost looks like a differential equation. To obtain a differential equation that we can use our standard tools on we choose a coordinate basis $|x\rangle$ and project our equation onto $|x\rangle$ via our inner product.

$$\langle x|\hat{D}|\psi\rangle = \langle x|(8\frac{d}{dx} + 5e^{\frac{d}{dx}})|\psi\rangle$$

$$= (8\frac{d}{dx} + 5e^{\frac{d}{dx}})\langle x|\psi\rangle$$

$$= (8\frac{d}{dx} + 5e^{\frac{d}{dx}})\psi(x)$$

$$= 8\frac{d\psi}{dx} + 5e^{\frac{d}{dx}}\psi$$

When we write explicit functions of some variable for a quantum state then that function is given by projecting the state vector onto that variable. viz

$$\langle x|\psi\rangle = \psi(x)$$

Technically, $|\psi\rangle$ is our vector and $\psi(x)$ is a little different. but for our purposes we think of the function $\psi(x)$ as a vector itself. QT involves seamlessly transitioning from states to functions and back to states and we think of them on the same footing but they are generally very different.

We will see when we make contact with reality that for some state $|\psi\rangle \in \mathcal{H}$ we can have MANY different functions that ALL represent the same quantum states. e.g.

$$\langle x|\psi\rangle = \psi(x)$$

$$\langle p|\psi\rangle = \widetilde{\psi}(p)$$

$$\langle E|\psi\rangle = \Psi(E)$$

Where all these different functions describe different aspects of our quantum state $|\psi\rangle \in \mathcal{H}$ and we will be able to transform from one to another. Also, note, the functional forms of these equations are all general vastly different. its not like $\psi(x) = x^2 \to \tilde{\psi}(p) = p^2$. No no no, $\psi \neq \tilde{\psi}$ even though they describe the same state.

When i imagine these things in my head i see $|\psi\rangle$ as this multi-faceted, high-dimensional, jagged, crystalline orb and you can view it from the x point of view, or the p point of view, or many many other points of view. the abstract vector encodes EVERYTHING about the system. It's amazing that nature winds up working this way.

Okay, so now we have to take AAALLLL the formulas we've previously engineered and apply it to functions such as $\psi(x)$. Yikes. It's not so bad.
Functional inner product:

$$\langle \phi|\psi\rangle = \int \phi^*(x)\psi(x)dx$$

We demand that $\psi(x)$ is such that:

$$|\psi(x)|^2 = \langle\psi|\psi\rangle = \int \langle\psi|x\rangle\langle x|\psi\rangle dx = \int \psi^*(x)\psi(x)dx < \infty$$

where we used the completeness property defined by

$$\int |x\rangle\langle x|dx = 1$$

In fact this completeness expression allows us to convert from discrete to continuous expressions with a push of a button. You start with your Hilbert Space vector $|\psi\rangle$ and multiply it by 1 accordingly. A lot of QT is really like building with legos: everything clicks together
Average of an observable:
We call the following expression the "average" of an observable $\hat{T}$

$$\langle\psi|\hat{T}|\psi\rangle = \int \langle\psi|x\rangle\langle x|\hat{T}|x\rangle\langle x|\psi\rangle = \int \psi^*(x)T\psi(x)dx$$

where

$$T = \langle x|\hat{T}|x\rangle$$

is the operator written in the x-basis. We will discover that expressions like these allow us to compute things like the average position or momentum or energy of a quantum system.

This concludes the main mathematical requirements we will need. Any other expressions or manipulations will occur in situ. We have finished the first part of our approach. A quantum system is described by a complex vector $|\psi\rangle$ that lives in a Hilbert space $\mathcal{H}$ with an inner product defined by

$$\langle\alpha|\beta\rangle = \alpha_1\beta_1 + \alpha_2^*\beta_2 + \ldots = \sum_{i=0}^{N} \alpha_i^*\beta_i$$

Such that, $\langle\alpha|\beta\rangle < \infty$

For our purposes this defines what we mean when we say "Hilbert Space"

Part 2:

We shall postulate what these states actually tell us. It is a postulate and nobody has given a complete proof for this interpretation but it is a natural extension of the historical route to quantum theory.

The Born Rule

Given a state and a basis below

$$|\psi\rangle = \sum_j \alpha_j|j\rangle$$

We will define a "measurement" to be a VIOLENT, absolute violent, projection of $|\psi\rangle$ onto some particular state related to an experiment that we can perform. Yes this all vague and nobody REALLY knows what's going on with this idea. Suppose we have a device that can measure our $|j\rangle$ states. The $|j\rangle$ are eigenvectors of some operator $\hat{J}$ and we have used those eigenvectors as our basis. When we perform measurement $\hat{J}$ we claim ONE and ONLY ONE eigenstate will be registered by our device. This single eigenstate will occur with some probability: we can't predict which state will be registered. Nobody knows. All that we know is that some state will crystallize into a definite form and this state $|j'\rangle$ will occur with a probability of $|\alpha_{j'}|^2$. for definiteness let's say $j' = 4$

$$|\psi\rangle \to |4\rangle$$

To find this probability we project $|\psi\rangle$ onto $|4\rangle$ by

$$\langle 4|\psi\rangle = \langle 4|(\sum_j \alpha_j|j\rangle) = \sum_j \alpha_j\langle 4|j\rangle$$

but $\langle 4|j\rangle = \delta_{4j}$

$$\sum_j \alpha_j \delta_{4j} = \alpha_1 0 + \alpha_2 0 + \alpha_3 0 + \alpha_4 1 + \cdots$$

Only the α_4 term survives this projection.

Define $p(n)$ as the probability of observing state $|n\rangle$

$$p(n) = |\langle n|\psi\rangle|^2 = \alpha_n^* \alpha_n = |\alpha_n|^2$$

This is our Born rule. It's the first tendril that we have reached out with to contact reality. It gives us the ability to make predictions. But how, exactly?

For a single quantum object this tells me NOTHIGN about which state i will observe when i perform an experiment. consider $\mathcal{H} = \mathbb{C}^2$. Our Hilbert space has two complex dimensions and therefore spanned by two complex vectors. Let's call our basis $|\uparrow\rangle$ and $|\downarrow\rangle$. If you like numbers rather than symbols call them $|0\rangle$ and $|1\rangle$ for all i care. If these are our basis vectors than any vector can be written as

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$$

That's it. That's the most general state. It's described by two complex numbers OR 4 real numbers.

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

When we measure our system we will ALWAYS either find it $|\uparrow\rangle$ or $|\downarrow\rangle$ but NEVER both! It's either a cat OR a dog, but not BOTH. Our QT does not tell us "before you measure it it is up". Your welcome to prepare a bunch of purely up states but if they interact with anything they might be destroyed. In a sense it is our measurements or observations or interactions that elucidate reality. Or, one takes the view that fundamental objects in reality are simply strange linear combinations of things that we can observe. It's up to you how you get your brain to internalize it but intuition only comes with exhaustive application of QT to realistic problems. How I conceptualize it changes from year to year.

So, we have this column vector $|\psi\rangle$ and if we run it through and up/down detector (called $\hat{S}$ for definiteness) we will only ever observe either $|\uparrow\rangle$ or $|\downarrow\rangle$ but NEVER both. It will be up $|\alpha|^2$ % of the time and down $|\beta|^2$ % of the time. By exhaustion we conclude that

$$\langle\psi|\psi\rangle = |\alpha|^2 + |\beta|^2 = 1$$

Our $\mathbb{C}^2$ space comes with a constraint: the total probability should equal 1. Four real numbers describing our state become just three. Also, notice that since we chose orthonormal eigenvectors to span our hilbert space we have

$$\langle\downarrow \mid \uparrow\rangle = \langle\uparrow \mid \downarrow\rangle = 0$$

This particular Hilbert space we've been inside is called a "Qubit". It's the foundation of quantum computing and we will work with it a LOT. It turns out that we could find ALL operators and observable for this simple system but we will postpone. What concerns us right now is the Born rule.

Okay, so a major theme of QT is asking the probabilistic question "what is the probability that _____

for each "blank" we ask we must build some kind of measurement apparatus that we can describe by a Hermitian matrix (or self-adjoint operator) which has an eigenvector that corresponds with that question.

Want to measure position? okay - a photographic plate or CCD might suit us. the we ask "what is the probability i find my object at position 5 on my photo?"

well, position is usually continuous so we form

$$\langle x = 5|\psi\rangle = \psi(5)$$

then the probability will be

$$|\langle x = 5|\psi\rangle|^2 = |\psi(5)|^2$$

So maybe there is a differential equation that allows us to solve, once and for, $\psi(x)$ that we could plug in any value and find the probability

The probability of an object being between position x and $x+dx$ is given by

$$p(x) = \psi^*(x)\psi(x)dx$$

and the total probability of being within some range is the integral over that range of this "probability density function"

$$\int_{universe} \psi^*(x)\psi(x)dx = 1$$

if we a priori know we have an object somewhere then it exists somewhere. the total probability of finding it somewhere in our lab or universe or what have you is 100%. This is what probabilities do.

Notice that $\psi(x)$ says that, generally, we have a linear combination over all possible positions. when we measure our object it might be here, or there, or in Andromeda with they're corresponding $\psi(x)$ values. Typically $\psi(x)$ is really big (it is a complex number, though) in a patch of the universe. Maybe I have an appreciable $\psi(x)$ only where a single hair on my head is (maybe i charged my hair with a balloon and i want to know what the chance is to observe a pack of charge at the tip of a hair). Maybe i want to know about where a cosmic microwave background photon who's wavefunction spans the entire universe is (yes, wavefunctions can be THAT big!!)! This is all what the function $\psi(x)$ gets us.

We call $\psi(x)$ "the wave function" and, confusingly, we call $\widetilde{\psi}(p)$ "the wave function". Sometimes quality folks will say "The Position Wave-Function" and "The Momentum Wave-Function". Recall that it is $|\psi\rangle$ that is our multifaceted, crystalline, continuously mutating, luminous, jagged, blob which describes the system. $\psi(x)$ and $\widetilde{\psi}(p)$ are merely single projections of this strange thing and the quantum system as a whole.

Now, one last thing that may be bugging you. I described this whole notion of "measurement" as a horrifically violent thing to do to such a beautiful luminous state blob. The reason i say it is so violent is because whatever information was encoded in the α and β complex numbers is forever lost to the universe. There is "no going back" after a measurement. It is also not a continuous procedure (as far as we know) and it seems to depend on whoever is interacting with our quantum system. Now, maybe these concerns are a bit too esoteric and minor at worst but later we will show that there is a continuous evolution of states, a rule they follow, a dynamics. It's a beautiful and pristine equation and whats even more awesome is that it's time symmetric. you can run the movie forward or backwards, none of the frames of the film go missing. Measurement is a second ugly evil cousin to our continuous evolution equation. Measurement destroys what may have been. Some folks even go so far as to say "measurement doesn't destroy anything, it creates a new universe for every possible outcome where a multi-verse exists and every possible outcome to every possible interaction and measurement is unfolding in parallel!" You can get a lot of interesting sci-fi culture out of that idea and it does solve our nasty measurement problem but ultimately it hasnt been shown to be testable and acts as an interpretations of quantum theory. There are many different interpretations of QT and they all have limitations. My preferred interpretation changes almost monthly. what can you do other

than learn about them all and imagine. Maybe you, dear reader, will develop a new interpretation or set to rest this measurement problem once and for all in a testable way via some clever insight!

This concludes our 2^{nd} task: The Born Rule. Next up are the specific operators that transform our states in Hilbert Space!

The algebra of operators

If we take hold of the Born rule and demand it (and for very good experimental reason we do!) then suddenly the types of matrices\operators that we're allowed to use to describe the transformations of quantum states become HIGHLY restricted. What would operators have to do with quantum states anyways? The Born rule was all about the states we can use so how could this have anything to do with operators or matrices.

We shall demand the following and then show that it must be true in order that the Born rule, and probability as a whole, is consistent.

Requirement: In order for probability to be valid any change in $|\psi\rangle$ must be induced by a unitary transformation.

We require $\langle\psi|\psi\rangle = 1$

Recall:

if $|\psi\rangle \in \mathcal{H}$ then $|\psi\rangle$ can be expanded as a linear combination of some basis states.

e.g.

$$|\psi\rangle = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix} = \sum_n c_n|n\rangle$$
$$= c_1|1\rangle + c_2|2\rangle + \dots$$

therefore,

$$\langle\psi|\psi\rangle = [c_1^* c_2^*, \dots.] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} = c_1^* c_1 + c_2^* c_2 + \dots$$
$$= \sum_{i=0}^{N} c_i^* c_i = \sum_{i=0}^{N} |c_i|^2 = 1$$

and since the $|c_n|^2$ represent the probability of obtaining state $|n\rangle$ then we require all the probabilities to sum to 1. i.e. "a 100% chance of the object

being in some state"

Suppose we operate on $|\psi\rangle$ with some operator $\mathscr{U}$. This $\mathscr{U}$ could be us manipulating the system in the laboratory or it could represent some interaction with something else in the universe far, far away from any lab. Whatever it is we expect the state to change somehow.

Then, abstractly, potentially changes our state to something other state. viz,

$$|\xi\rangle = \mathscr{U}|\psi\rangle$$

but we still require $\langle\xi|\xi\rangle = 1$. Whatever complicated state the object is in it can still be expanded in a set of basis states and all the probabilities must still sum to 1. Reasonable, no? So we require (recall that the row vector is the adjoint of the vector and dagger switches the order!)

$$1 = \langle\xi|\xi\rangle = \langle\psi|\mathscr{U}^\dagger\mathscr{U}|\psi\rangle$$

$$\Rightarrow \mathscr{U}^\dagger\mathscr{U} = \mathbb{1}$$

Whatever happens in this quantum world those $\mathscr{U}$ operations are VERY special. There are TONS of matrices that dont behave this way, yet in quantum mechanics our operations on states MUST satisfy $\mathscr{U}^\dagger\mathscr{U} = \mathbb{1}$ if we take the Born rule as an axiom of QM. I wouldnt know how to make sense of anything otherwise! Recall, the Born rule says that $|c_i|^2$ are to be interpreted as probabilities.

Matrices that satisfy $\mathscr{U}^\dagger\mathscr{U} = \mathbb{1}$ are called "Unitary".

$$\mathscr{U}^{-1} = \mathscr{U}^\dagger$$

Example,

$$\mathscr{U} = e^{i\phi}$$

is a 1×1 unitary matrix/operator because

$$\mathscr{U}^\dagger = e^{-i\phi} \Rightarrow \mathscr{U}^\dagger\mathscr{U} = e^{-i\phi}e^{i\phi} = 1 \quad \forall\phi \in \mathbb{R} \quad QED$$

Easy Peasy!

We can have 2 by 2 unitary matrices. For example,

$$\mathscr{U} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\mathscr{U}^\dagger = \mathscr{U}^{-1}$$

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_{2\times 2}$$

Okay, this is actually a Hermitian matrix (called $\hat{\sigma_y}$) but Hermitian matrices can be unitary too. This would be a type of observable (indeed an important one!).

Another example,

Scattering

Suppose we fire an e^-at something. Initially $e^-$is in state $|\psi\rangle$ after it scatters or interacts it's in a new state $|\psi'\rangle$

$$\therefore |\psi'\rangle = \hat{S}|\psi\rangle$$

where S is unitary.

S is called "the S-matrix" and encodes all the complicated details involved in scattering. It is in general VERY difficult to solve for and an entire industry of mathematical and physical methods exists solely for determining S! (if you've read ahead please dont confuse this S with spin S, different things, same notation).

Finding all unitary matrices and observables for a Qubit.

Consider the Hilbert space $\mathcal{H} = \mathbb{C}^2$. Our precious and important Qubit state (this could be atom with two primary energy levels such as the Yb^{3+}, it could be a little magnet, a mirror cavity, a superconducting junction, and much much more. There are MANY representations of a qubit)

The structure of these unitary matrices can generally be seen as follows. We have a 2-complex dimensional Hilbert space and we choose some basis for this space. ONLY AFTER we pick a basis can we begin writing the elements of the $\mathscr{U}$ matrix! So, in some basis of our choosing we can write some $\mathscr{U}$ down that looks like:

$$\mathscr{U} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

$$\mathscr{U}^\dagger = \begin{bmatrix} \alpha^* & \gamma^* \\ \beta^* & \delta^* \end{bmatrix}$$

$$\mathscr{U}^\dagger \mathscr{U} = \mathbb{1}$$

Dear reader, please dont get these α and β and such mixed up with our previous notation for up and down. This is a brand new discussion so the two

assignments don't necessarily match. these greek letters are each a complex number and describe the matrix $\mathscr{U}$. Now, because $\mathscr{U}^\dagger\mathscr{U} = \mathbb{1}$ we can find relations among these greek guys by multiplying the matrices and setting them equal to the identity (DO IT!). you will find

$$\Rightarrow \gamma\alpha^* + \delta\beta^\star = 0 \Rightarrow \delta = \frac{-\alpha^*}{\beta^*}$$

$$\alpha\alpha^* + \beta\beta^* = 1 \Rightarrow \gamma = -\beta^*$$

$$\Rightarrow \mathscr{U} = \begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix}$$

There we have it. Every operation that can ever be done on a qubit, that preserves the Born rule is right there.

We will also demand that $\alpha^2 + \beta^2 = 1$ so that $\det(\mathscr{U}) = 1$ but this cant completely be explained at the moment. For now please trust it's a reasonable thing to require.

Now, in our qubit Hilbert space $\mathcal{H} = \mathbb{C}^2$ any state can be written in any basis we please. Often we pick our favorite observable and then describe a state in terms of the eigenvectors of that observable.

Let's choose a particular Hermitian matrix σ_z as our special observable. We dont yet have any physical reason to choose this particular observable yet but we will find later that this is a VERY important physical observable and for now it's the barebones simplest we could possibly choose. It is,

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

In the basis of it's own eigenvalues. We start by finding the eigenvalues, there should be two. Since we have a diagonal matrix the eigenvalues are simply the diagonal elements but heck, let's solve it anyways

$$\begin{aligned} \det\left(\sigma_z - \lambda\mathbb{1}\right) &= \det\begin{bmatrix} 1-\lambda & 0 \\ 0 & -1-\lambda \end{bmatrix} = 0 \\ &= (1-\lambda)(-1-\lambda) = 0 \\ &\Rightarrow \lambda = \pm 1 \end{aligned}$$

Great, we can find eigenvalues of anything that way! Now, the eigenstates are defined by

$$\sigma_z|i\rangle = \lambda_i|i\rangle$$

Show that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are the eigenvectors. Well, this is kind of "duh" because we chose a diagonal matrix! This is what choosing a diagonal matrix means - its eigenvectors are those above!

We will call $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$ just to use a different notation than $|\uparrow\rangle$ and $|\downarrow\rangle$ which is pretty typical for these things. They're just names

Now that we have a basis we can write any possible state as a linear combination of the definite states $|0\rangle$ and $|1\rangle$

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle$$

We call this a "super-position" of definite states $|0\rangle$ and $|1\rangle$.

We need 2 complex numbers c_0, c_1 to specify a specific state AND since we're doing quantum mechanics we require $c_0^2 + c_1^2 = 1$ (total probability equals 1). This means we need just three $\mathbb{R}$ numbers to specify our state. We can visualize these states as points on a sphere called the "Bloch Sphere"

So any unitary transformation merely moves $|\psi\rangle$ around on this sphere and measurements force $|\psi\rangle$ to either definite state $|0\rangle$ or $|1\rangle$ with probabilities c_0^2, c_1^2. This is the entirety of a single qubit system!

Measuring the definite states $|0\rangle$ and $|1\rangle$.

Now that we picked a basis we can measure whether $|\psi\rangle$ is in $|0\rangle$ or $|1\rangle$. We did all this already and if

$\langle 0|0\rangle = 1$ and $\langle 0|1\rangle = 0$

that is to say the basis states are orthonormal then

$$\therefore \quad \text{prob}(|\psi\rangle \to |0\rangle) = c_0^2$$

A measurement like this can be performed by sending electrons through a special magnet. This is the so-called Stern-Gerlach experiment that led to the discovery that electrons are well described as these two-state systems.

Okay, that's all of the unitary operations possible on a qubit. What are all of the possible observables?

Recall that in QM observables are self-adjoint or Hermitian matrices

$$M^\dagger = M$$

$\mathbb{1}$ is certainly self adjoint. Consider an arbitrary M whose elements are complex numbers.

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow M^\dagger = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}$$

if $M = M^\dagger$ then

$$a = a^*, b = c^*, c = b^*, d = d^*$$

$$\therefore M = \begin{bmatrix} a & b \\ b^* & d \end{bmatrix} = c_0\mathbb{1} + c_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$
$$+ c_3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

So we find that $M = c_0\mathbb{1} + c_1\sigma_x + c_2\sigma_y + c_3\sigma_z$ has a very particular form. We "choose" to write M this way, we could have chosen a completely different way to decompose the general matrix M. Often we consider matrices themselves as being "vectors" which we can write as a linear combination of "basis matrices"

$\sigma_x, \sigma_y, \sigma_z$ defined above are called the Pauli Matrices and comprise a VERY important basis for observables in our qubit system. We will find that they also generate a group called $SU(2)$. The special unitary group in two dimensions. $SU(2)$ are sometimes called "generalized rotations" but we wont yet go into that.

Let's note some properties of the Pauli matrices:

$$\sigma_i^2 = \mathbb{1}$$

σ_i is like the square root of $\mathbb{1}$. That's kind of neat i guess.

$$[\sigma_i, \sigma_j] = \sigma_i\sigma_j - \sigma_j\sigma_i = 2i\varepsilon_{ijk}\sigma_k$$

i.e. $[\sigma_x, \sigma_y] = 2i\sigma_3$ and so forth.

$[\sigma_i, \sigma_j]$: This bracket mapping is extremely important not just in physics but mathematics as a whole. It is called "the commutator" or "Lie Bracket".

$[,] : M \times M \to M$

This of this operation like you would addition or division, or any other basic operation. Addition is a mapping that takes two numbers and spits out a number. We dont often write it like tyhis but in functional form + is a function on the reals

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

or

$$\cdot : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$$

is the vector dot product. it takes two vectors and produces a number.

Our $[,]$ mapping takes two Hermitian matrices as inputs and returns a Hermitian matrix. It is enormously important that $[\sigma_i, \sigma_j]$ always returns the other basis matrix σ_k just like adding two natural numbers always returns another natural number. Doing this mapping you can never leave the set you began with. This feature is called an "abstract algebra" or simply "algebra". This commutator is incredibly crucial for QT.

The commutator essential tells us sort of how changing the order of operators acts. For example, you're used to the dot product of two vectors not making a difference whether $a \cdot b$ or $b \cdot a$

$$[a, b]_{dot} = 0$$

for the dot product only.

We could (and should!) define the anti-commutator, which is the same but with a plus sign, but we wont for now. Anyhow these are simply important added structure that we mention in passing. We will have much more to say on them later.

So there we have it; all the possible observations we could do on a qubit. Any linear combination of $\mathbb{1}, \sigma_x, \sigma_y, \sigma_z$ comprise an observable of a qubit or two level system and we very often choose eigenstates to be those of our σ_z observable.

Now, the qubit is the smallest non-trivial system we could build and we can completely determine everything but typically $\mathcal{H}$ is BIG. And I mean REALLY big. 100 qubits have 2^100 (~10^{30})possible states! We typically treat electrons as qubits when ignoring their positions and simply cant solve such a measly number. Imagine how big $\mathcal{H}$ might be with a Mol of electrons all zipping about and bumping into each other! No chance! We simply dont have the computational power or mathematical tools to analyze such a space without making some MAJOR approximations. We can write down a model that describes ALL macroscopic matter (e.g. a table or something) but we have no hope in solving them. A lot, however, can be discerned, discovered, and observed from these approximations as we will see.

Part 4: Imposing the symmetries of the universe onto our mathematical framework.

Our universe seem to have some very nice basic properties. We've come to notice that in our very narrow existence that when we set up an experiment

in Milan and measure reality and then travel to Tokyo and perform the same experiment they seem to agree quite accurately despite MANY things being different. Milan and Tokyo are two different places in our universe and, what's more, we're doing experiments at two different points in time. Evidently our universe evolved, things moved about, in the time between our experiments. Well, you say, the changes of the rest of the universe are locally very small here on earth. E.g. as we travel around the sun in a slightly elliptic path the various forces on the earth due to the sun, the moon, Jupiter, etc are all slightly different and we could potentially account for those slight changes in our two same experiments at two different places and times in the universe and we're pretty satisfied by that.

Or are we? What if the very LAWS of physics were dependent on space and time themselves? What if $F = ma$ here on Earth but $F = qV/x + ma$ three light years away? What if the laws of physics themselves depend on where and when you are in our universe?! Couldnt that be a possibility? Maybe conservation of energy hold in our area but doesn't somewhere else? We certainly don't know for sure. However, ALL(?) of the observations humanity has made here and as we peer into the heavens implies that the laws of physics should be universal. Wherever and whenever you are $F = ma$ modulo technical details.

CLAIM:
Our universe is such that translation of spatial and time coordinates is a symmetry. i.e. if we take our universe and shift everything to the right by, I don't know, 50 lightyears, then everything will behave exactly the same. Same thing with time: if we freeze the universe somehow and then unfreeze it 45 billion years from now (whatever it even means to have time outside of a frozen universe!) the evolution should be identical to if we didn't freeze it. Furthermore, we could rotate the entire universe by 33^o and we expect the evolution to be the same.

Dear reader, some things are so obvious that we never consider them deeply yet once we do question the things we consider obvious they become quite mysterious! This is one of those things and when we consider it deeply and work out all the consequences we find beautiful and deep connections between apparently unrelated ideas in physics.

One could say a MAJOR contribution Einstein made to humanity was along these lines: he asked, "what are the consequences of demanding various symmetries?". E.g. he held something called "Lorentz symmetry" (from the study of electricity and magnetism) to be a symmetry of our universe and from that germ of an idea he developed special and general relativity!

This symmetry approach to physics is an entirely new way to consider

nature and the mathematics that models it. the past 120 years has been all about "what are the symmetries of the system i am studying?"

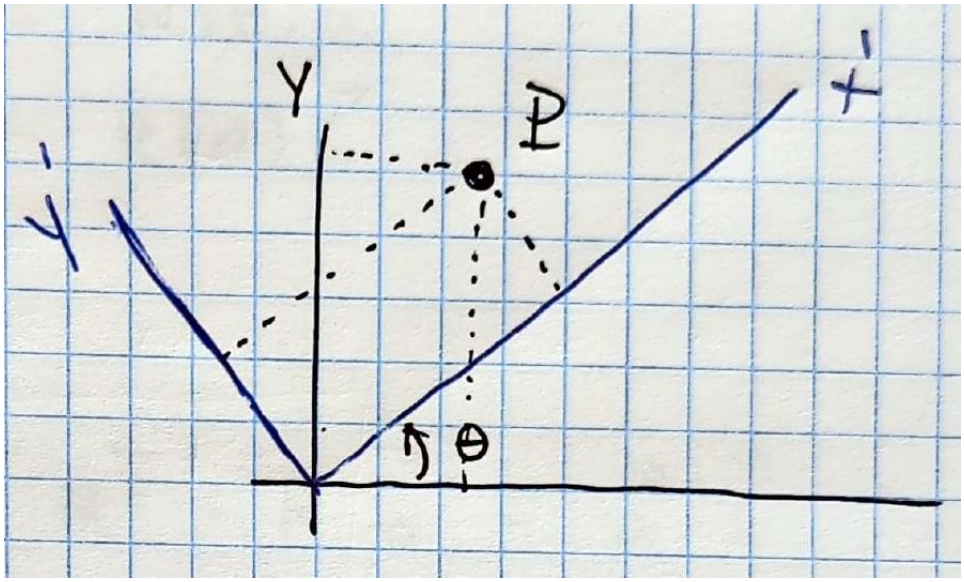
We shall develop a new mathematical tool that is based on symmetry that we shall leverage to further elucidate this mysterious $|\psi\rangle$ and how it is required to evolve in a universe respecting the symmetries that we wish to demand. You could play the same game with OTHER symmetries we know the universe doesn't demand and you can have a lot of fun building different quantum physics that simply dont exist and you can study those universes as a hobby or as deep research. We will impose these basic symmetries and will see what this QT has to say as a consequence of those symmetries.

The mathematical field that studies symmetry is called "Group Theory". Eugene Wigner was the first physicist to apply Group Theory to physics with MAJOR success. However, oddly, for 40 years he was essentially the only one. Most physicists of those 40 years called Group Theory, "gruppen-pest". It was weird and esoteric to them. It was only around the $1950-1960$'s that "gruppen-pest" started to catch fire. Hiesenberg, Murray Gell-Mann, Abdus Salam, and others led the path forward by discovering symmetries in the zoological collection of hundreds upon hundred of sub-atomic particles. They found certain underlying symmetries that had gone completely unnoticed.

We will learn a bit about what they did much later but for now we have more meager aims - what kind of behavior is exhibited by a $|\psi\rangle$ in a universe with time-translation, spatial-translation, and rotational symmetries? We are taking our mathematical framework and now reaching out to the universe with our tendrils and seeing if we can make contact with anything realistic. If this framework can make predictions that are testable, then we are happy campers (even happier if the predictions matches our experimental results!). If these mathematics doesnt make any predictions or if it makes testable predictions that assuredly dont match reality then we fail and try again. Again, this is not the historical route towards discovering quantum physics here on Earth but perhaps the way it was discovered far far away in some galaxy on some planet by some kind of alien physicists. Its fun to ponder other paths to the same things we've learned here on Earth. Almost always discovery is messy and the beautiful theory is slowly pieced together through trial and error. Here we begin with beauty and check whether nature is described by it. It's the much more difficult route towards discovery because beauty does NOT imply reality!

Let's study symmetry in the world we're most familiar with and see if this symmetry point of view leads us down the correct path.
Suppose you and I have rotated coordinate systems as below and share an

origin.



For some point P we have the relationship between our two coordinate systems

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

We can write this as a matrix equation

$$\vec{x'} = R_\theta \vec{x}$$

where R_θ is a matrix operator given by

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

behold we can rotate vectors at will! consider two vectors $\vec{a}$ and $\vec{b}$

$$\vec{a}\cdot\vec{b} = \text{ Some number}$$

$\vec{a}\cdot\vec{b}$ we expect should be "invariant" under these kinds of rotations above. It is completely independent of the coordinates we choose.

Suppose,

$$
\begin{aligned}
&\vec{a}' = R_\theta \vec{a} \\
&\vec{b}' = R_\theta \vec{b} \\
&\vec{a}' \cdot \vec{b}' = \vec{a} \cdot \vec{b} \\
\therefore \quad &\vec{a}' \cdot \vec{b}' = a^T R_\theta^T R_\theta \vec{b} = \vec{a} \cdot \vec{b}
\end{aligned}
$$

Notice that when we take the dot product of two vectors the first object must always be a row vector. $\vec{a}$ is technically a column vector and so we convert it to a row vector by using the transpose operation. $\vec{a}^T$ is a row vector, i.e. $\vec{a}^T = [a_x \quad a_y]$

Now, the above manipulations imply that

$$
\Rightarrow \quad R_\theta^T R_\theta = \mathbb{1} \Rightarrow \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
$$

Hooray! It works! We aren't done yet with the classical rotations but to go further we must develop a specific tool in mathematics called Lie Group Theory. We now leave our classical physics and discuss symmetry from a mathematical point of view. I will use quantum notation for vectors and matrices and such and then we may adapt the notation for classical physics considerations as well as quantum (notational differences are just notation and culture!).

Lie Groups

Suppose you have a family of (unitary) operators that can be described by a continuous variable s (for example, rotation matrices depend on the angle θ). The matrices don't HAVE to be unitary outside of our quantum theory. For classical rotations we will find non-unitary matrices - rotations in (x, y, z) coordinates that dont necessarily have complex numbers in the matrices.

Define $\mathscr{U}(0) = \mathbb{1}$ as the identity matrix. The identity is at parameter $s = 0$ and $\mathbb{1}$ is certainly unitary.

Furthermore, suppose we demand that

$$
\mathscr{U}(s_1 + s_2) = \mathscr{U}(s_1)\,\mathscr{U}(s_2)
$$

This type of structure is called a "homo-morphism" and is the biggest part of defining a mathematical object called a "Group". There are other requirements of a Group but theyre pretty obvious (e.g. there's an inverse, an identity, etc). Prove that a 2D vector rotation satisfies this group homomorphism.

Now, consider $s \ll 1$. Since s is a continuous parameter then we can Taylor expand our operators around $s = 0$ as follows (we will switch from script $\mathscr{U}$ to capital Latin U for notational clarity)

$$U(s) = U(0) + \left.\frac{dU}{ds}\right|_{s=0} ds + \cdots\cdots$$

for $ds = s - 0 \ll 1 \quad ds^2 \approx 0$ so we drop all higher terms. This is a Taylor expansion applied to continuous matrices.
Since $U(s)$ is unitary we require

$$U^\dagger U = \mathbb{1}$$

Multiply this out using our Taylor approximation for $U(s)$ and find

$$\mathbb{1} + ds \left[\frac{dU}{ds} + \frac{dU^\dagger}{ds}\right]\Bigg|_{s=0} + O\left(s^2\right) = \mathbb{1}$$

The only way to satisfy this unitary requirement is if, for infinitessimal s we have

$$\frac{dU}{ds} = -\frac{dU^\dagger}{ds}$$

call $\frac{dU}{ds} = iK$ then K must he Hermitian

$$(iK)^\dagger = -iK^\dagger \Rightarrow K = K^\dagger$$

We call this matrix/operator K is called "the generator" of the U transformation.

If you know K then you know everything about $U(s)$! There is a sense in which this K is the "tangent vector of the Group itself, near the identity". Lie Groups, themselves, are generally curved "manifolds". We wont define further but there are some groups that are manifolds/geometries like spheres. We will study one later which is equivalent to the 3-sphere S^3 - a sphere embedded in 4 dimensions!

Notice the magic!

$$\begin{aligned} U(s_1 + s_2) &= U(s_1)\, U(s_2) \\ \frac{dU(s_1 + s_2)}{ds_2} &= U(s_1)\, \frac{d}{ds_2} U(s_2) \\ &= U(s_1)\, iK \end{aligned}$$

Now take $s_1 = 0$ and notice that this is a differential equation whose formal solution is

$$U(s) = e^{iKs}$$

How do we exponentiate a matrix? We leverage Taylor yet again to define the exponential of a matrix

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots$$
$$e^{iKs} = 1 + iKs + \frac{1}{2}(iKs)^2 + \cdots$$

This is essentially building up the full $U(s)$ from the generator matrix K which amounts to infinitely many tiny infinitesimal transformations.

i.e. if you want to rotate 120° you can do so 1^0 at a time 120 times. Same idea!

So, we have found that if we have a group of transformations parameterized by a continuous variable s then we can write ANY unitary matrix in this group as

$$\mathscr{U}(s) = U(s) = e^{iKs}$$

and this looks just like how we represent complex numbers a la Euler. Notice, in particular that the Hermiticity of K allows

$$\mathscr{U}^\dagger \mathscr{U} = e^{iKs}(e^{iKs})^\dagger = e^{iKs}e^{-iK^\dagger s} = e^{iKs}e^{-iKs} = \mathbb{1}$$

which proves this $\mathscr{U}(s)$ we found is assuredly Unitary!.

This method will be VERY handy in quantum mechanics and many other fields of science! If you know the generators of symmetry then you know almost everything (there are subtleties we will discuss soon).

Example of Lie's Method in Classical Physics

Now lets see if we could do the same thing with rotations as we did with our general unitary matrices.

Finding the generators of Rotations

We play the same game: a little rotation at a time. We notice when $\theta = 0$ $R_{\theta=0} = \mathbb{1}$ so we will Taylor expand this rotation matrix around $\theta = 0$

We have,

$$R(\theta) \approx \mathbb{1} + A \quad \text{for } \theta \ll 1$$
$$R^T R = \mathbb{1} = \left(\mathbb{1} + A^T\right)(\mathbb{1} + A) \approx \mathbb{1} + A^T + A$$

note: θ is "inside" A and A^TA is of order θ^2 so we neglect these higher terms for small rotation.

Thus we require $A^T = -A$ in order that $R^TR = \mathbb{1}$. This A is an anti-symmetric matrix.

For 2 dimensional spaces there is only one fundamental anti-symmetric matrix: We will factor out θ from our A and write $A = J\theta$ where

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \Rightarrow A = \theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

So for small θ

$$R \approx \mathbb{1} + \theta J = \begin{bmatrix} 1 & \theta \\ -\theta & 1 \end{bmatrix}$$

But what if we want to do an arbitrary rotation rather than an infinitessimal rotation? For arbitrary θ we can divide θ up into N pieces and then peform MANY MANY tiny rotations N times and then take the limit as $N \to \infty$
viz,

$$\begin{aligned} R_\theta &= \lim_{N\to\infty} \left[R\left(\frac{\theta}{N}\right) \right]^N \\ &= \lim_{N\to\infty} \left[\mathbb{1} + \frac{\theta J}{N} \right]^N = e^{\theta J} \end{aligned}$$

$$e^{J\theta} = \mathbb{1} + J\theta + \frac{J^2\theta^2}{2!} + \cdots$$

Try and show that each entry of this expansion reproduces the sins and cosines of our original representation of R_θ! Notice, this rotation is not

We call the rotations in $2D$ space the group $SO(2)$. This terminology comes from the requirements on how the rotation matrices behave.

$SO(2)$ literally means "the special orthogonal group in 2 dimensions"

"special" means $\det R = 1$
"orthogonal" means $R^TR = \mathbb{1}$

Assignment: Try and show $SO(3)$ is generated by three matrices J_x, J_y, J_z for three dimensional vectors $\vec{r}$

$$\vec{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \vec{a} \cdot \vec{b}' = \vec{a} \cdot \vec{b}$$

You will find

$$J_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad J_y = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

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

and a general rotation is given by a vector of angles $\theta_x, \theta_y, \theta_z$

$$R(\vec{\theta}) = e^{\theta_x J_x + \theta_y J_y + \theta_z J_z} = e^{\Sigma \theta_i J_i} = e^{\vec{\theta} \cdot \vec{J}}$$

Continuous groups like rotations and the unitaries found above are called "Lie Groups"

The generators of the Lie Group form what's called a "Lie Algebra"

The Lie algebra for $SO(3)$ is given by

$$[J_i, J_j] = J_i J_j - J_j J_i = i\varepsilon_{ijk} J_k$$

where

$$\varepsilon_{ijk} = -\varepsilon_{jik}$$

is the totally antisymmetric symbol

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{for} \quad 123, 231, 312 \\ -1 & \text{for} \quad 213, 132, 321 \\ 0 & \text{for} \quad \text{all others} \end{cases}$$

e.g.

$$[J_x, J_y] = iJ_z$$
$$[J_y, J_z] = iJ_x$$
$$[J_z, J_x] = iJ_y$$

Almost everything interesting about $R(\vec{\theta})$ is contained in this Lie algebra! We often think in terms of $[J_i, J_j]$ instead of $R(\theta)$ directly. The matrices themselves we say "form a 3-dimensional representation of $SO(3)$"

But we are NOT required to use matrices explicitly. The Lie Algebra is more general than matrices. it says "any abstract object that behaves as $[J_i, J_j] = i\varepsilon_{ijk} J_k$ is THE generator of $SO(3)$"". They don't have to be in a finite dimensional basis, they could be infinite dimensional operators.

Another "representation" of $SO(3)$ is given by

$$J_z = -i\left[x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right]$$

$$\begin{aligned}[J_x, J_y] &= \left[-i\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), -i\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\right] \\ &= iJ_z\end{aligned}$$

These are all differential operators and act on objects spanning an infinite dimensional space of functions. No matrices in sight yet the differential operations satisfy the Lie Algebra of $SO(3)$! You, the astute reader, may even recognize these particular combinations of derivatives: they are angular momentum operators! We find that angular momentum operators are elements of the Lie Algebra of $SO(3)$ and generate rotations. How cute!

Again, ANY representation that satisfies this specific Lie Algebra generates $SO(3)$ via exponentiation! This gives meaning to weird things like $e^{\frac{d}{dx}}$.

$$e^{a\partial_x} f(x) = f(x+a)$$

This expression yields the Taylor expansion. We will find soon that this expression is closely related to momentum!

Symmetry in Quantum Systems

If we know operations in our Hilbert space must be unitary and we know our universe is rotationally/translationally symmetric, then if we can "represent" a rotation by a unitary operator then the observable properties of the state $|\psi\rangle$ should remain unchanged by these special operations.

i.e. if A is some observable (which operationally means we can build a detector to measure such a property: for example, position) with eigenstates $|\phi_n\rangle$ then

$$\hat{A}\,|\phi_n\rangle = a_n\,|\phi_n\rangle$$

This is what it means to make an "observation", say $+1$ for the state $|\uparrow\rangle$ or some other state whose eigenvalues we are able to observe. That's just what the above equation says.

Now, after some (say, coordinate) transformation we require the observables (eigenvalues!) remain unchanged,

$$\hat{A}' \left|\phi'_n\right\rangle = a_n \left|\phi'_n\right\rangle$$

No prime on the a_n eigenvalues - they are equal in both frames (watch the primes like a hawk here! primes are the transformed objects relative to our original coordinates).

This expression means that you and I must agree on observable quantities. It would be absolutely amazing if you measured a cat and saw a dog!

Essentially, if you transform the states (via a coordinate change) then the operators/observables must also transform in a way that leaves the eigenvalues of those operators unchanged. The eigenvalues are the values that we can measure and record. Sometimes, we're a bit casual and say things like "I measure the state". We measure eigenvalues - those are real quantities. States are abstract and complex valued beasts. Our eigenvalues correspond to states but there doesnt need to be any reason why those abstract states are precisely the same in different coordinates - we just need to agree on their eigenvalues!

So, the state vectors may look different and the forms of the operators may look different, but the measurable quantities must be the same (unless you physically do something to the system!). Observables values are coordinate independent. Therefore, consider the transformation,

$$\hat{A} \left|\phi_n\right\rangle \longrightarrow \hat{A}' \left|\phi'_n\right\rangle$$

but $\left|\phi'_n\right\rangle = \mathscr{U} \left|\phi_n\right\rangle$ for some $\mathscr{U}$

$$\begin{aligned} &\therefore \tilde{A}' \mathscr{U} \left|\phi_n\right\rangle = a_n \mathscr{U} \left|\phi_n\right\rangle \\ &\Rightarrow \mathscr{U}^{-1} \tilde{A}' \mathscr{U} \left|\phi_n\right\rangle = a_n \left|\phi_n\right\rangle \end{aligned}$$

Remember to solve a matrix equation like above we multiply both sides on the LEFT with $\mathscr{U}^{-1}$. We have shown that states transform as

$$\left|\psi'\right\rangle = \mathscr{U} \left|\psi\right\rangle$$

and operators transform as

$$\hat{A}' \rightarrow \mathscr{U}^{-1} \hat{A} \mathscr{U}$$

These formula will be VERY important for our further exposition. Say it again with gusto: "states transform with a $\mathscr{U}$ and operators are sandwhiched by $\mathscr{U}^{-1}$ and $\mathscr{U}$"!

Now, let's consider our very first full quantum symmetry: time-translation.

Time-translation symmetry:

Let us label our quantum state to be dependent on time. This is relatively reasonable because we suspect quantum states in our universe can evolve in time somehow and there must be some unitary operator governing that evolution for probabilities to sum to 1. We write now, $|\psi(t)\rangle$ to describe our state at different points in time and operators are static. We could, however take a different approach where states are static and operators evolves and change in time. These are called the "Schrödinger" and "Heisenberg" representations respectively. We are considering the "Schrödinger" picture and you may wish to prove that they are entirely equivalent!

Since we model time as a continuous parameter t then we can apply our Lie Group machinery to it and find the generators of our time translation symmetry!

Consider a quantum state that depends on time.

$$|\psi(t)\rangle \in \mathcal{H}$$

let $\hat{T}(a)$ be a unitary operator that acts on $|\psi\rangle$ as

$$\hat{T}(a)|\psi(t)\rangle = |\psi(t+a)\rangle$$

i.e. it moves\translates our state forward in time by amount a

Since we demand unitarity then $\hat{T}$ can also also move $|\psi\rangle$ backwards in time.

$$\hat{T}^{\dagger}(a) = \hat{T}^{-1}(a)$$
$$\hat{T}^{-1}(a)|\psi(t)\rangle = |\psi(t-a)\rangle$$

Following Lie's method we shall Taylor expand $\hat{T}(a)$ around the identity $\mathbb{1}$

Since a is taken to be infinitesimally small in Lie's approach, let's go ahead and simply call it dt since we're among friends here.

then

$$\hat{T}(dt) = \hat{T}(0) + \left.\frac{d\hat{T}}{dt}\right|_{t=0} dt + \cdots\cdots$$

impose $\hat{T}^{\dagger}T = \mathbb{1}$ (since $\hat{T}$ is unitary) and find

$$\left(\mathbb{1} + \frac{dT^\dagger}{dt}dt\right)\left(\mathbb{1} + \frac{dT}{dt}dt\right) = \mathbb{1}$$
$$\Rightarrow \quad \frac{dT^\dagger}{dt} = -\frac{dT}{dt}$$

since $dt^2 \approx 0$. Let's call $\frac{dT}{dt} = \frac{i\hat{H}}{\hbar}$ (it's convenient to include the i and a physical constant $\hbar$ where, for notational convenience we shall set $\hbar = 1$ and reintroduce it into our final equations). Then ,

$$(i\hat{H})^\dagger = -(i\hat{H})$$

Implies that

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

So that $\hat{H}$ is Hermitian.

Back to the defining equation

$$\hat{T}(dt)|\psi(t)\rangle = (\mathbb{1} + i\hat{H}dt)|\psi\rangle$$

$$= |\psi(t)\rangle + i\hat{H}dt|\psi(t)\rangle = |\psi(t+dt)\rangle$$

rearranging and recalling that dt is infinitessimal we find a derivative pops out in the limit $dt \to 0$!

$$\frac{|\psi(t+dt)\rangle - |\psi(t)\rangle}{dt} = \frac{i}{\hbar}\hat{H}|\psi(t)\rangle$$
$$= \frac{d}{dt}|\psi(t)\rangle = \frac{i}{\hbar}\hat{H}|\psi(t)\rangle$$

This is a complicated version of a very standard differential equation. It is of the form $f' = af$ which has exponential solutions via the method of "separation of variables"

$$\Rightarrow |\psi(t)\rangle = e^{\frac{i}{\hbar}Ht}|\psi(0)\rangle$$

Now, since the units of $\left[\frac{d}{dt}\right] = \frac{1}{\text{second}}$ and we demand the exponential to be unit-less. This is why we went ahead and introduced $\hbar$. It's the only constant lying around that causes the dimensionalful quantities to work out right. Namely,

$$\left[\frac{1}{dt}\right] = [\frac{H}{\hbar}]$$
$$[\hbar] = J \cdot s$$

$$\Rightarrow \left[\frac{H}{\hbar}\right] = \frac{\text{Something}}{J \cdot s}$$

Since we are doing quantum theory, a physical description of nature, $\hbar$ has to show up somewhere in our equations and we find that it is yoked to $\hat{H}$. We have,

$$\Rightarrow [H] = J$$

is an energy!

So we have $|\psi(t)\rangle = e^{\frac{i\hat{H}t}{\hbar}}|\psi(0)\rangle$ and we say " $\hat{H}$ generates time-translations'. If you want to move forwards or backwards in time then you can do so via $\hat{H}$.

$U(t;t') = e^{\frac{i\hat{H}(t-t')}{\hbar}}$ is called the "propagator" and governs the evolution of a quantum state in time.

Now, since we expect, for constant energy, time-translation to be a symmetry, then if we choose as the states we wish to observe as eigenstates of the $\hat{H}$ operator, then we will expect these measurable quantities (energy) to be constant in time. It seems natural to choose as a basis for $\langle$ to be the eigenstates of $\hat{H}$ since these wont have any changing eigenvalues (because thats what we often want to measure).

We dont ALWAYS want to measure energy. If we want to measure some other property, like angular momentum, then we might want to leverage some other operator whose eigenstates are angular momentum states. Symmetries make it easier because multiple experimenters can agree on the measured quantities. e.g. if energy wasnt constant in time then theres no reason to expect energy states to persist. If I have a box of hornets and i measure the energy and then you come in and shake up the box of hornets and measure the energy then of course they would be in different states and our observations (eigenvalues) would generally disagree. So we save ourselves the trouble and look for states that are eigenstates of an operator which generates a symmetry.

We will do this for the Hydrogen atom later. Find the symmetries and describe the states in terms of the generators, that way when I look at a Hydrogen atom today from some place in the universe and you look at it tomorrow at some other place we can agree that we're both looking at a Hydrogen atom in the same quantum state!

Emphatically "EIGENVECTORS OF GENERATORS OF SYMMETRIES MAKE GREAT STATES WITH WHICH TO FORM A BASIS OF HILBERT SPACE"

OKAY, SO If we solve the eigenvalue/eigenvector equation for $\hat{H}$ then we will have a simple expression for how $|\psi(t)\rangle$ evolves. Choosing the energy basis we solve the eigenvalue/eigenstate equation

$$\hat{H}\,|n\rangle = E_n\,|n\rangle$$

and

$$\frac{d}{dt}|\psi(t)\rangle = \frac{i}{\hbar}\hat{H}|\psi(t)\rangle$$

we call these $|n\rangle$ the "energy states". You, or others, might want to write them as $|E_n\rangle$. It's your choice how to symbolize it. If the energy of the system is neither increasing nor decreasing then the states are constant in time. What CAN change are the probability amplitudes of each $|n\rangle$ in a given expansion of $|\psi\rangle$

We have thus derived the single most important equation in the entire study of quantum theory.

The eigenvalue equation above is the "Time-Independent Schrödinger Equation" and the equation with the time-derivative is known as the "Time Dependent Schrodinger Equation"(usually written with a partial differential when spatial coordinates are considered)! We dont know yet how to represent $\hat{H}$ specifically but let's briefly recap how we got to this momentous accomplishment.

1. We postulated reality is based on complex vectors in a Hilbert Space
2. We can't predict WHAT exactly will happen but we CAN find probabilities for any possible experimental result.
3. The universe is symmetric under time translation and systems whose energy remains constant are invariant under time translation
4. Following Lie's advice we generated time translations by $\hat{H}$
5. Finally, the eigenstates of the generator of time translations are energy states and the eigenvalues are invariant under transformation of time.

Dear reader, we have done VERY little physics here. Again, this is not how Schrödinger's equation was discovered but it could have been. Historically it was VERY messy and complicated to get to where we are.

Digression aside, we can solve this equation to obtain the energy states and energy levels (eigenvalues) and then if we would like to determine how our general state evolves in time we can hit it with the propagator!

viz,

$$\begin{aligned}|\psi(t)\rangle &= e^{\frac{i}{\hbar}Ht}\,|n\rangle \\ &= e^{\frac{i}{\hbar}E_n t}\,|n\rangle\end{aligned}$$

The $e^{\frac{i}{\hbar}E_n t}$ factor "looks like" this state is changing which we said "no no no. observables (energy) are constant in time!"" But dear reader, states are complex valued and abstract. Notice that

$$\begin{aligned}\langle\psi(t)|\psi(t)\rangle &= \langle n|\, e^{-\frac{iE_n t}{\hbar}} e^{\frac{iE_n t}{\hbar}}\,|E_n\rangle \\ &= \langle n|n\rangle = 1\end{aligned}$$

and so we say that these $|n\rangle$ states are "stationary". Their "phase" is rotating but our observations are always the same. The exponential term in this energy basis amounts to only a rotating phase and if we are strictly interested in a specific state $|n\rangle$ then the phases matter not to our observations.

However, the phases are SUPER important if $|\psi\rangle$ is a linear combination of energy states. Then the phases can lead to all sorts of time-dependent interference and such and probabilities might change in time. The eigenvalues dont change, but the probabilities of obtaining a particular eigenvalue can certainly change. viz

$$|\psi\rangle = \sum \alpha_n |n\rangle$$

then,

$$U(t;t')|\psi\rangle = U(t;t')\sum \alpha_n|n\rangle$$

$$= \sum \alpha_n U(t;t')|n\rangle$$

$$= \sum \alpha_n e^{\frac{i}{\hbar}\hat{H}t}|n\rangle$$

$$= \sum \alpha_n e^{\frac{i}{\hbar}E_n t}|n\rangle$$

Note that in third to fourth line we converted $\hat{H}$ to E_n apropos the energy basis states $|n\rangle$

This is more or less how we will solve the Hydrogen atom. The math will be much more complicated but the procedure will more or less be the

same. We shall leverage symmetries in order to write the simplest possible time-independent basis states.

Time Translation Symmetry and Momentum

We expect physics to be invariant under time-translation. You've seen this in your very first introduction to physics where it's emphasized that YOU get to choose whatever coordinates YOU want but ultimately YOU and I should be able to predict and measure the same things. You throw a rock off a mountain and you measure its speed and time to impact of the ground to be whatever. Then I, watching from the bottom, measure the same values (modulo the time it takes light to travel from the top to my measuring equipment).

What's more is we learn that coordinate system in motion relative to each other at constant speed is also a symmetry called Galilean. Einstein, Lorentz, and Poincare took this further and extended Gallilean symmetry to the so-called "lorentz symmetry" of special relativity. We will discuss Lorentz and Galileo later but for now let us focus on the space-translational symmetry of our universe.

Everything follows entirely the same way as the case of time-translation (in fact we couldve done them together using a 4-vector of space and time coordinates, say $x = (t, x, y, z)$). Instead, we shall focus here only on spatial coordinates in one dimension.

Essentially we need only change symbols in our equations. States depend now on position variables as well as time.

$$|\psi(x)\rangle \in \mathcal{H}$$

let $\hat{T}(a)$ be a unitary operator that acts on $|\psi\rangle$ as

$$\hat{T}(a)|\psi(x)\rangle = |\psi(x+a)\rangle$$

i.e. it moves\translates our state forward in space by amount a

Since we demand unitarity then $\hat{T}$ can also also move $|\psi\rangle$ backwards in space.

$$\hat{T}^{\dagger}(a) = \hat{T}^{-1}(a)$$
$$\hat{T}^{-1}(a)|\psi(x)\rangle = |\psi(x-a)\rangle$$

Following Lie's method we shall Taylor expand $\hat{T}(a)$ around the identity $\mathbb{1}$

Since a is taken to be infinitesimally small in Lie's approach, let's go ahead and simply call it dt since we're among friends here.

then

$$\hat{T}(dx) = \hat{T}(0) + \left.\frac{d\hat{T}}{dx}\right|_{x=0} dx + \cdots\cdots$$

impose $\hat{T}^\dagger T = \mathbb{1}$ (since $\hat{T}$ is unitary) and find

$$\left(\mathbb{1} + \frac{dT^\dagger}{dx}dx\right)\left(\mathbb{1} + \frac{dT}{dx}dx\right) = \mathbb{1}$$
$$\Rightarrow \quad \frac{dT^\dagger}{dx} = -\frac{dT}{dx}$$

since $dx^2 \approx 0$. Let's call $\frac{dT}{dx} = \frac{i\hat{p}}{\hbar}$ (it's convenient to include the i and a physical constant $\hbar$ where, for notational convenience we shall set $\hbar = 1$ and reintroduce it into our final equations). Then ,

$$(i\hat{p})^\dagger = -(i\hat{p})$$

Implies that

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

So that $\hat{p}$ is Hermitian.

Back to the defining equation

$$\hat{T}(dx)|\psi(x)\rangle = (\mathbb{1} + i\hat{p}dx)|\psi\rangle$$

$$= |\psi(x)\rangle + i\hat{p}dx|\psi(x)\rangle = |\psi(x+dx)\rangle$$

rearranging and recalling that dt is infinitessimal we find a derivative pops out in the limit $dt \to 0$!

$$\frac{|\psi(x+dt)\rangle - |\psi(x)\rangle}{dx} = \frac{i}{\hbar}\hat{p}|\psi(x)\rangle$$
$$= \frac{d}{dx}|\psi(x)\rangle = \frac{i}{\hbar}\hat{p}|\psi(x)\rangle$$

This is a complicated version of a very standard differential equation. It is of the form $f' = af$ which has exponential solutions via the method of "separation of variables"

$$\Rightarrow |\psi(x)\rangle = e^{\frac{i}{\hbar}\hat{p}x}|\psi(0)\rangle$$

Now, since the units of $\left[\frac{d}{dx}\right] = \frac{1}{\text{meters}}$ and we demand the exponential to be unit-less. This is why we went ahead and introduced $\hbar$. It's the only constant lying around that causes the dimension-ful quantities to work out right. Namely,

$$\left[\frac{1}{dx}\right] = [\frac{p}{\hbar}]$$
$$[\hbar] = J \cdot s$$
$$\Rightarrow \left[\frac{p}{\hbar}\right] = \frac{\text{Something}}{J \cdot s}$$

Since we are doing quantum theory, a physical description of nature, $\hbar$ has to show up somewhere in our equations and we find that it is yoked to $\hat{p}$. We have,

$$\Rightarrow [p] = kg\frac{m}{s}\frac{s}{s}\frac{m}{m} = J\frac{s}{m}$$

is a momentum and causes our exponential to be dimensionless!

So we have $|\psi(x)\rangle = e^{\frac{i\hat{p}x}{\hbar}}|\psi(0)\rangle$ and we say " $\hat{p}$ generates spatial-translations'. If you want to move forwards or backwards in space then you can do so via $\hat{p}$.

We have found that momentum generates spatial translations and energy (or more correctly the "Hamiltonian") generates translations in time. So that if you wish to move an object into a future position and time then these operators tell you how to proceed!

I like to hope that these two examples have felt rather logical and well founded. If youre at all confused you can apply the same analysis to classical vectors. Make a change of coordinates and study how things like dot-products depend on these continuous changes (i.e. dot products we hope should be invariant!). You can apply Lie's method to these translations as well (its the same idea as we did with classical rotations).

Representation Theory of $SO(3)$

Now we press onward towards rotational symmetry in quantum systems. To do this we shall talk about the mathematical theory of "Representations".

A Group is a set of transformations that leave some object invariant. Mathematically it has a more precise definition but not by much. Some groups you can keep in mind are translations, rotations, the real numbers

under addition, flipping triangles, and so forth. What we REALLY care about is representing a "Group" in terms of matrices or operators. a mathematical group is "more abstract" than the matrices we typically use to describe it. I was confused by this for many years so I hope to alleviate that same confusion in you, dear reader.

We define a representation of a group G to be such that for all members $g \in G$ there exists a corresponding matrix $D(g)$ such that if

$$g_1 \circ g_2 = g_3$$

then

$$D\left(g_1\right) D\left(g_2\right) = D\left(g_1 g_2\right) = D\left(g_3\right)$$

this is called a group "Homo-morphism". Practically speaking it just means we can "represent" our abstract group G as a set of matrices that "act like" the group on some kind of vector space. CRUCIALLY: IT DOES NOT MATTER WHAT VECTOR SPACE so long as it works. it make work for vectors, it may work for tensors, it may work for functions, it may work for elephants (if they could form a vector sace) If $G = SO(3)$ A set of matrices that does the job are what we typically think of as our everyday rotation matrices acting on a vector space composed of columns (x, y, z)

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

These are 3×3 matrices and therefore "act" on 3×1 column vectors. This is A vector space which satisfies the requirements of $SO(3)$ GROUP. There could be others like 5×1 columns, or 7654×1 columns matters not so long as $SO(3)$ can be represented. Viz,

$$\vec{x}' = R\vec{x}$$

just like we are used to.

Again, emphatically, this "representation" of 3×3 matrices is just ONE CHOICE (OR possibility). We might be able to find $4 \times 4, 5 \times 5, 30 \times 30,$

et cetera sized matrices that obey the "rules" of $SO(3)$ and indeed there are MANY different representations of our $SO(3)$ group..

Recall the "rules" of $SO(3)$ were

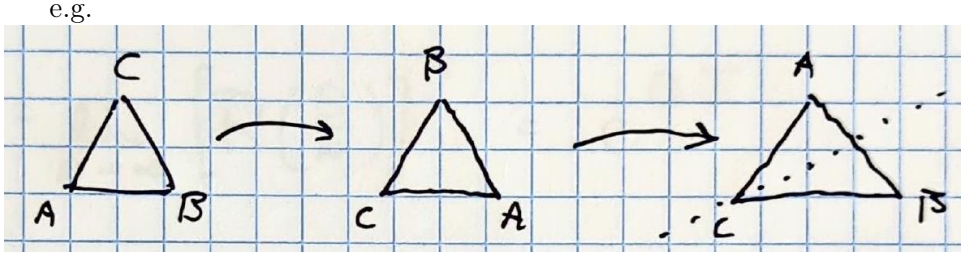
1. "special" $\Rightarrow \det R = 1$
2. "orthogonal" $\Rightarrow R^\top R = \mathbb{1}$
3. $R_i(\theta_1) R_i(\theta_2) = R_i(\theta_1 + \theta_2)$ (group homomorphism)

Note that 3) is only for a single direction (e.g. x, y, or z). $R_x(\theta_1) R_z(\theta_2) \neq R_x(\theta_1 + \theta_2)$ or something. We will see how to compose directions later.

Now $SO(3)$ just happens to be a group with continuous parameters (e.g. $\theta_x, \theta_y, \theta_z \in \mathbb{R}$). Not all groups are continuously parameterized.

These types of groups are called "Lie Groups". Some groups are discrete, for example, the symmetries of an equilateral triangle form a "Finite Group". Nothing continuous about it; just flip the triangle around such that it looks the same (the vertices need to be the same).

e.g.



So, Lie groups can be parameterized by continuous variables and this makes them EXTREMELY interesting and useful. For example, EVERY finite group can be shown to be equivalent to a simple permutation of elements. That's kind of boring. Lie Groups are rich with structure and nuance. We have our group representation matrices as follows:

$$R_x(\theta), R_x(\theta + \delta\theta), \ldots \text{ et cetera}$$

there are uncountably many different θ 's since $0 \leq \theta \leq 2\pi \in \mathbb{R}$.

Now previously we Taylor expanded a general $SO(3)$ matrix near the identity $\mathbb{1}$ and enforced the "special" and "orthogonal" conditions on our matrices

$$R(\theta) \approx \mathbb{1} + J\theta + \cdots$$

and $\theta \ll 1$ so we ignore terms of order θ^2; above

$$R^T R = \mathbb{1} \quad \Rightarrow \quad J^T = -J$$

so that a "small" rotation is given by

$$R \approx 1 + \theta J = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix} \quad (+ \text{ higher terms } \approx \theta^2 \approx 0)$$

We get a full rotation by an arbitrary angle by doing lots of little rotations.
The full Rotation by θ is

$$R(\theta) = \lim_{N\to\infty} \left[R\left(\frac{\theta}{N}\right) \right]^N = e^{\theta J}$$

for $SO(2)$ we found

$$e^{\theta J} = \cos\theta \mathbb{1} + \sin\theta J = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in SO(2)$
Back to $SO(3)$ we have 3 matrices J_x, J_y, J_z

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad J_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad J_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and $R(\vec{\theta}) = e^{\theta_x J_x + \theta_y J_y + \theta_z J_z} = e^{\vec{\theta}\cdot\vec{J}}$
where $\vec{\theta}$ is a set of three angles $\theta_x, \theta_y, \theta_z$
In physics we often "pull out a $i = \sqrt{-1}$ " and write

$$R(\vec{\theta}) = e^{i\vec{\theta}\cdot\vec{\mathcal{J}}}$$

where $\mathcal{J}_x = iJ_x \quad \mathcal{J}_y = -iJ_y, \quad \mathcal{J}_z = -iJ_z$
Henceforth, we will write J_x, J_y, J_z instead of the script $\mathcal{J}$'s
e.g.
J_x is now $= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$ et cetera

Now we find our $SO(3)$ representation matrices are UNITARY!

$$R(\vec{\theta}) = e^{i\vec{\theta}\cdot\vec{\mathcal{J}}}$$

This is now a rotation we could potentially apply to a quantum state $|\psi\rangle$!!
What a nice little trick. No Hilbert space insight yet we began with classical

rotations and we found the quantum rotations. In hindsight it couldn't have been any other way because we see that macroscopic manifestations of quantum objects and we rotate these with our everyday rotations. But this leaves open a very subtle but crucial opening: we may have missed some "rotational-type" representations. Keep this in mind as we go further into $SO(3)$. We will next find ALL possible (unitary) representations of $SO(3)$. We will find there are infinitely many! Prepare yourself with the most difficult and abstract analysis yet! Dont worry about understanding why we would go this route, just that once we start this route, i hope that each step is logically clear and justified.

The Lie Algebra

These three J's tell us almost everything we need to know about the Group $SO(3)$

They behave as

$$[J_i, J_j] = J_i J_j - J_j J_i = i\varepsilon_{ijk} J_k \tag{La}$$

e.g $[J_x, J_y] = iJ_z$

ANY matrix satisfying this "Lie Algebra" is a valid representation of a generator of $SO(3)$. We say that these J_i form the "Lie Algebra" $\mathfrak{so}(3)$. The algebra is different than the full group. The algebra is the "tangent space" of the group near the identity (e.g. the tangent plane of a sphere for example). Lie groups have VERY interesting geometries. $SO(3)$ in particular almost looks like a 3-sphere but antipodal points are "glued together"(Zee offers a lucid discussion in his Group Theory in a Nutshell textbook). An easier group to visualize is $SO(2)$. This group has the geometry of a circle. The tangent space at $\theta = 0$ is a one dimensional line so the Lie Algebra should have only one element (which was what we discovered).

Now, the question becomes: "How can I find all the possible "representations" of $SO(3)$? Do any $N \times N$ matrices satisfy the rules of $SO(3)$ besides the standard x,y,z coordinates or just some?"

It turns out that only $(2j+1) \times (2j+1)$ matrices satisfy the Lie Algebra $\mathfrak{so}(3)$ if $j \in \mathbb{N}$ $j = 0, 1, 2, 3, \ldots$ etc. For any j there will always be 3 generators of $\mathfrak{so}(3)$

$$J_x, J_y, J_z \to R(\theta) = e^{i\vec{J}\cdot\vec{\theta}}$$

We have already found the $j = 1$ representation matrices of $SO(3)$. Explicitly, $j = 1 \Rightarrow (2j + 1) \times (2j + 1) = 3 \times 3$ matrices: the J_i matrices above.

The $j = 1$ representation has many names such as "the defining irrep of $SO(3)$" or "the vector irrep of $SO(3)$ or the "triplet" of $SO(3)$ and further still, the "Spin-1" irrep of $SO(3)$. These ALL mean the same thing. Sometimes people will say such and such "furnish" the representation. Imprecises and confusing terminology abound! We shall usually say "Spin-j" when we specify a particular representation. Finally, apologies in advance but it is typical to refer to BOTH $R(\theta) \in SO(3)$ and $J_i \in \mathfrak{so}(3)$ as a spin-j representation. The literature is confusing that way.

Moving on, how could we possibly fid the $j = 5$ set of three 11×11 matrices that obey/represent $SO(3)$?! What on earth would even be "rotating" that's 11-dimensional?! e.g.

For $j = 1$ we easily imagine vectors $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and

$$\vec{x}' = R\vec{x}$$

but I cant imagine any 11-dimensional object of our universe (or mathematics) that rotates somehow under such a matrix.

Note: Importantly these irreps may have NOTHING to do with space or time

In fact, there ARE spin-2 objects of spatial coordinates that DO obey $SO(3)$. The rotation matrices are 5×5 and we call the objects which those matrices act on "tensors". For example, a tensor can be represented as a matrix with elements

$$M_{ij}$$

Where rotation matrices "act on" one index at a time. File this away in your "curious but not needed for now" drawer.

Constructing the objects that $(2j + 1) \times (2j + 1)$ matrices of $SO(3)$ act on is an algorithmic task but often we need not compute such things directly and instead rely on a higher level abstract notation.

We shall find all possible representations by employing the following algorithm. This algorithm will also work for $SU(2)$, $SO(N)$, or $SU(N)$ or any other Lie groups we may be interested in.

We perform the following steps.

1. Complexify the Lie Algebra
   e.g.

$$J_+ = J_x + iJ_y$$

$$J_- = J_x - iJ_y$$

2. Choose the eigenvectors of J_z. i.e. $J_z|m\rangle = m|m\rangle$
3. Ask "what is $J_z J_+|m\rangle$" and "what is $J_z J_-|m\rangle$"
4. Use $[J_i, J_j] = i\varepsilon_{ijk}J_k$ and rewrite it in terms of J_+, J_-, J_z Work this out! e.g.

$$[J_z, J_\pm] = \pm J_\pm$$
$$[J_+, J_-] = 2J_z$$

5. Use these commutation relations for step 3
6. Use the fact that eigenvectors are orthogonal. i.e. $\langle n|m\rangle = \delta_{mn}$ to find any undetermined coefficients.

It's complicated and weird that anyone discovered this but if we follow our noses and follow the procedure we will find all the eigenvectors of J_z in any $j \in \mathbb{N}$ representation! In particular, the number of eigenvectors we will find in this process will span the entire vector space of our spin-j representation.

In this algorithm we usually choose J_z to be diagonal.

e.g. for $j = 1$

$$J_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

is not diagonal but if we happen to have diagonal matrices it gets much easier. Regardless we CAN find eigenvectors for non-diagonal matrices and then we would re-write J_z in diagonal form such that the eigenvalues for each eigenvector will be the diagonal elements (we do this in detail below).

This procedure has another MAJOR benefit for physics: the eigenvectors of the generators or the symmetry group are the states that we observe!

We don't so much care about the matrices as much as we care about the states themselves! The matrices kind of tell us which states are best for a basis. That is precisely what were going to find from this method! The states

of angular momentum from the group of rotations will be useful states to study because angular momentum is often conserved. These states will span our particular Hilbert space and any general state $|\psi\rangle$ could be expressed as a linear combination of these angular momentum states. That's our whole motivation.
Let's begin

$$J_z|m\rangle = m|m\rangle$$

This is entirely general. I don't know if J_z is $3\times 3, 8\times 8, 5000\times 5000$. No matter how big or small it is, it WILL have eigenvectors and eigenvalues: 3, 8, and 5000 of them respectively. m labels which eigenvector within the total set of them. e.g. if $j = 1$ we will choose to label $m = -1,\ 0,\ 1$. the value of m is called "the weight": math lingo

Step 3 - Find

$$J_z J_+|m\rangle$$

from $[J_z, J_\pm] = \pm J_\pm$we have

$$\begin{aligned} J_z J_+ - J_+ J_z &= J_+ \\ \therefore\quad J_z J_+ &= J_+ + J_+ J_z \end{aligned}$$

Now we plug in what we just found from the commutation relation above:

$$\begin{aligned} J_z J_+|m\rangle &= (J_+ + J_+ J_z)\,|m\rangle \\ &= J_+|m\rangle + J_+ J_z|m\rangle \end{aligned}$$

but we know $J_z|m\rangle = m|m\rangle$! So we can evaluate

$$\begin{aligned} J_z J_+|m\rangle &= (J_+ + mJ_+)|\,m\rangle \\ &= (m+1)J_+|m\rangle \end{aligned}$$

What does this tell us?

It says "when you apply J_z to the state $J_+|m\rangle$ you get $(m+1)J_+|m\rangle$" call $J_+|m\rangle$ some other name, say $|q\rangle$. It's just some kind of column vector. Let's call it specifically $|m+1\rangle$
then

$$J_z J_+|m\rangle = (m+1)J_+|m\rangle$$

says that

$$J_z|m+1\rangle = (m+1)|m+1\rangle$$

But this is our eigenvector/eigenvalue equation for J_z! This says $J_+|m\rangle$ is an eigenvector of J_z with eigenvalue $(m+1)$!

So it seems that J_+ moves us from one eigenvector of J_z to another eigenvector. In particular,

$$J_+|m\rangle = c_{m+1}|m+1\rangle$$

We dont know what c_{m+1} is yet but we can worry about that detail later....it's just some number that depends on m. The state is most important.

Similarly, in exactly the same way,show that

$$J_-|m\rangle = b_{m-1}|m-1\rangle$$

Now if J_z is a $(2j+1)\times(2j+1)$ matrix then there will be $2j+1$ eigenvectors and we expect there will be a "lowest" and a "highest" "weight-m" vector such that

$$J_+|m_{\text{max}}\rangle = 0$$
$$J_-|m_{\text{lowest}}\rangle = 0$$

If we're at the highest weight state, then if we try and go any higher we will fall off the ladder. We get zero. Similarly, for the lowest we will step off the ladder.

Let us call

$$m_{\text{max}} = j \quad \text{and} \quad m_{\text{lowest}} = -j$$

It's just notation, j labels the eigenstates. Then we have a set of $2j+1$ states arranged just like a ladder. The bottom rung is $m=-j$ and the top rung is $m=j$ for the spin-j representation.

That's it!! We have just found ALL states of ALL representations of $SO(3)$! They are the $(2j+1)$ states labelled by $j \in \mathbb{N}$.

We say "the spin-j representation of $SO(3)$ is spanned by:"

$$|-j\rangle, |-j+1\rangle \ldots |j-1\rangle, |j\rangle$$

For example,

$$j = 0 \Rightarrow 1 \times 1 \text{ matrices } \Rightarrow |0\rangle$$
$$j = 1 \Rightarrow 3 \times 3 \text{ matrices } \Rightarrow |-1\rangle, |0\rangle, |1\rangle$$
$$j = 2 \Rightarrow 5 \times 5 \text{ matrices } \Rightarrow |-2\rangle, |-1\rangle, |0\rangle, |1\rangle, |2\rangle$$

and so on! For completeness I merely cite what c_{m+1} and b_{m-1} are. You can use the stated facts (e.g. the eigenvectors are orthonormal) to find them.

$$\begin{aligned} J_+|m\rangle &= c_{m+1}|m+1\rangle \\ J_-|m\rangle &= b_{m-1}|m-1\rangle \\ c_{m+1} &= \sqrt{(j+m+1)(j-m)} \\ b_{m-1} &= \sqrt{(j+m-1)(j+m)} \end{aligned}$$

In QM these are the states of "orbital" angular momentum! In chemistry class you learned about the $s, p, d, f, \ldots$ orbitals. Those were secretly these $SO(3)$ representations!
$j = 0$ is an s-orbital
$j = 1$ is a p-orbital
$j = z$ is a d-orbital
and so on.

We often write $\ell = 0, 1, \ldots$ instead of j when referring ONLY to orbital angular momenta. j typically refers to "total" angular momentum. We will see later why we make this distinction. Here allow me to switch to ℓ

In QM the generators of $SO(3)$ are the orbital angular momentum operators.

$$L_x, L_y, L_z$$

for $l = 0 = s$-states $L_x, L_y, L_z = 0$
for $l = 1 = p$-states they are

$$\hat{L}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \hat{L}_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad \hat{L}_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

They act on states $|-1\rangle, |0\rangle, |1\rangle$. When you want to measure orbital angular momentum in a particular direction you apply the operator in that direction.

We are always free to transform these $\hat{L}_x, \hat{L}_y, \hat{L}_z$ into a different forms using a unitary transformation.

e.g. There exists some U such that

$$\mathscr{U}^\dagger \hat{L}_z \mathscr{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \tilde{L}_z$$

$$\mathscr{U}^\dagger \hat{L}_x \mathscr{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \tilde{L}_x$$

$$\mathscr{U}^\dagger \hat{L}_y \mathscr{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} = \tilde{L}_y$$

We only transform these because in this form L_z is diagonal. In particular the eigenstates match up with what we did above in the general case.

$$|-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad |0\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad |1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

For $l = 2$ we have three 5×5 matrices.

$$L_z = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

I have NO idea what L_x & L_y are for $1 = 2$ but L_z can always be constructed diagonally from the eigenvalues of the states in the way we found above.

$$|-2\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad |-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \cdots$$

Finally, no matter what representation you have of $SO(3)$ a rotation about some angle $\vec{\theta}$ is always given by

$$R(\theta_x, \theta_y, \theta_z) = e^{i(\theta_x L_x + \theta_y L_y + \theta_z L_z)}$$

In particular there are other ways to describe a spatial rotation. When you use a screwdriver you certainly don't say "I will rotate by $\theta_x, \theta_y, \theta_z$!"Yikes! That'd be difficult to get around in life. No, we typically define a direction $\vec{n}$ and then rotate about that direction by some angle. E.g. imagine using a screw driver. We will explicitly write our rotations this way later on.

Notice that this R is unitary by construction! We usually dont care much about $R(\theta_x, \theta_y, \theta_z)$. In QM almost everything is all about the eigenvectors of the generators of a symmetry group-

In this case

$$[L_i, L_j] = i\varepsilon_{ijk}L_k$$

Also, super important!
these

$$|j\rangle, |j-1\rangle, \ldots |-j\rangle$$

states are NOT NOT NOT eigenvectors of L_x or L_y
e.g. $L_x|m\rangle \neq m|m\rangle$

L_x on state $|m\rangle$ world be some crazy combination of states, e.g. for $l = 1$ rep

$$L_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$L_x|-1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$== \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}}|0\rangle$$

$$\therefore \quad L_x|-1\rangle = \frac{1}{\sqrt{2}}|0\rangle$$

which is definitely not an eigenvector of L_x

So this is how we chose to construct everything. We usually measure angular momentum with respect to the z-axis of whatever our coordinate system is and if we're interested in other directions we simply measure those others.

There is, however, one super-important operator (called a Casimir operator) that does indeed share eigenvectors with L_z. It is

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

and acts on state $|m\rangle$ as

$$L^2|m\rangle = \ell(\ell+1)|m\rangle$$

Use $(L\pm, L_z)$ to show this

Classically $\sqrt{L^2} = \sqrt{L_x^2 + L_y^2 + L_z^2}$ "looks like" the length of the angular momentum vector.

e.g.

$$\ell = 0 \Rightarrow \sqrt{\ell(\ell+1)} = 1 \qquad \ell = 1 \Rightarrow \sqrt{\ell(\ell+1)} = \sqrt{1}$$
$$\ell = 2 \Rightarrow \sqrt{\ell(\ell+1)} = \sqrt{6}$$

and so forth.

In QM the length of the z-component of angular momentum is less than the total length!

This is unusual because suppose we tried to align $\vec{L}$ perfectly with our z-axis; in that case we expect the magnitude of L_z to be equal to $\sqrt{L^2}$ However, in QM the Lie Algebra says "No, No, No".

$$\sqrt{L^2}|m\rangle = \sqrt{\ell(\ell+1)}|m\rangle$$
$$L_z|m\rangle = m|m\rangle$$

and m is always less then $\sqrt{\ell(\ell+1)}$

e.g. $\ell = 2$ and $|m\rangle = |2\rangle$ is our highest state then $\quad 2 < \sqrt{6}$

Evidently if we know the value of L_z then we automatically do NOT Know L_x or L_y! This is ultimately because of

$$[L_i, L_j] = i\varepsilon_{ijk}L_k$$

in physics L_i comes with an $\hbar$ so

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k$$

if $\hbar \to 0$ then

$$[L_i, L_j] = 0$$

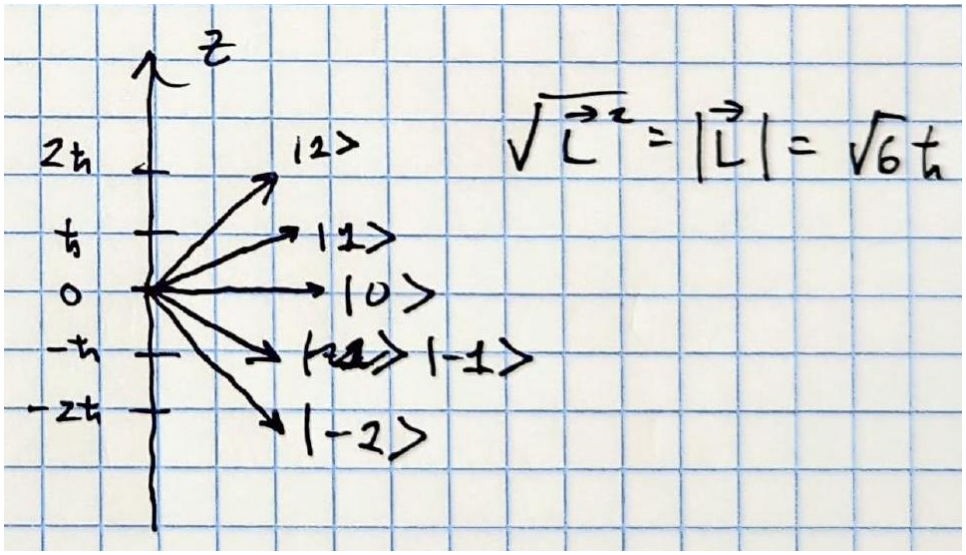
This is classical physics land with no quantum around. We are allowed to know L_x, L_y, L_z exactly with no problems in classical physics. In QM we are restricted a la Heisenberg (more to say on this later).

Quantum theory is a bit more weird than our sensibilities. We only know the magnitude and z-component of orbital angular momentum and we live with the uncertainty in L_x and L_y

In the previous $\ell = 2$ case we have $\sqrt{\ell(\ell+1)} = \sqrt{6}$

For state $|0\rangle$ the angular momentum lies in a circle in the $x, y,$ plane given by

$$L_x^2 + L_y^2 = 6$$



We have an entire circle of possible values for L_x and L_y when $L_z = 0$ for this 0 state of L_z

Keep in mind NONE of this includes "spin" in case you were wondering. We will next learn of another type of angular momentum that $SO(3)$ can't inform us of: the intrinsic angular momentum of a quantum object called "spin". We see no Pauli matrices anywhere.....yet!

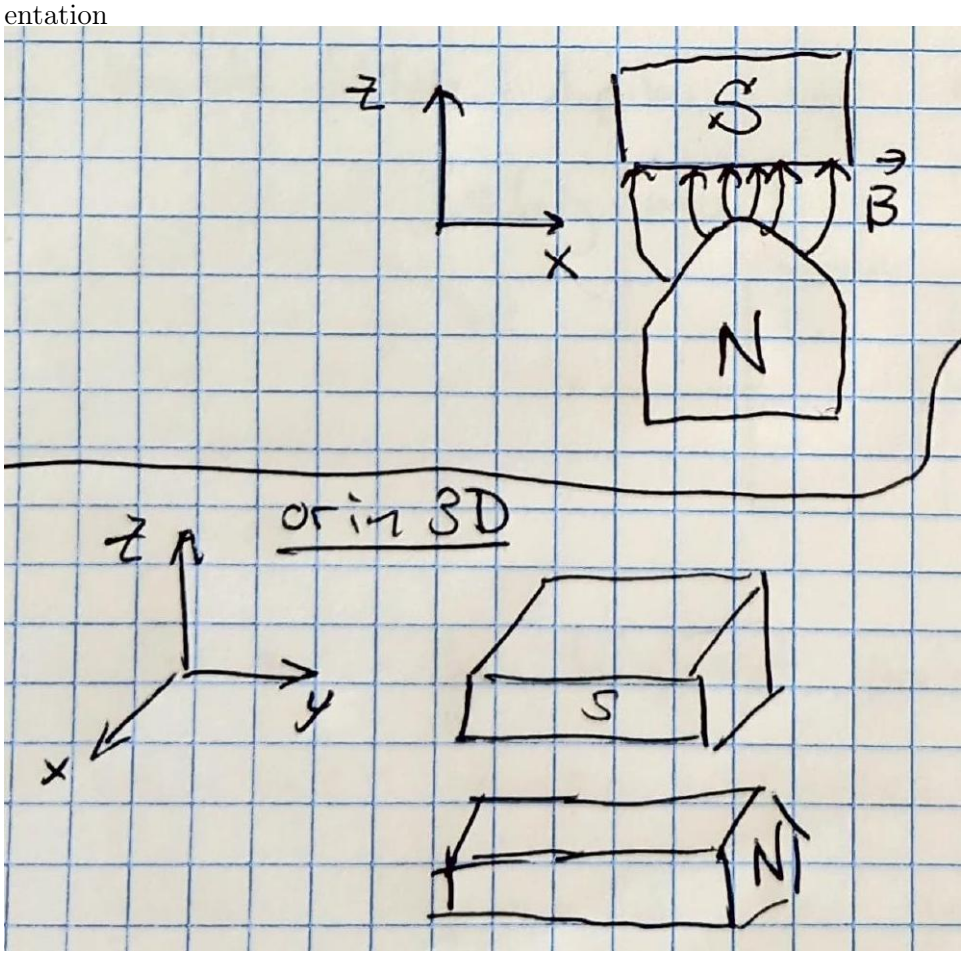
Internal Symmetry and Spin

The Stern-Gerlach experiment

In 1921 and 1922 Otto Stern and Walter Gerlach proposed and conducted the following experiment by asking the question "What happens when we

send silver atoms through a region of space containing an inhomogeneous magnetic field?"

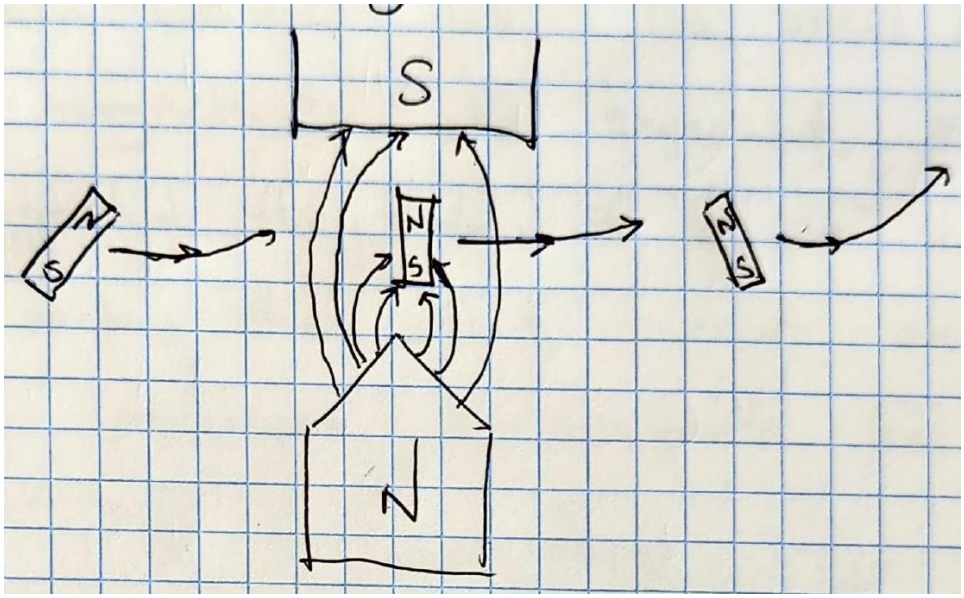
The magnetic field was created using two magnets in the following orientation



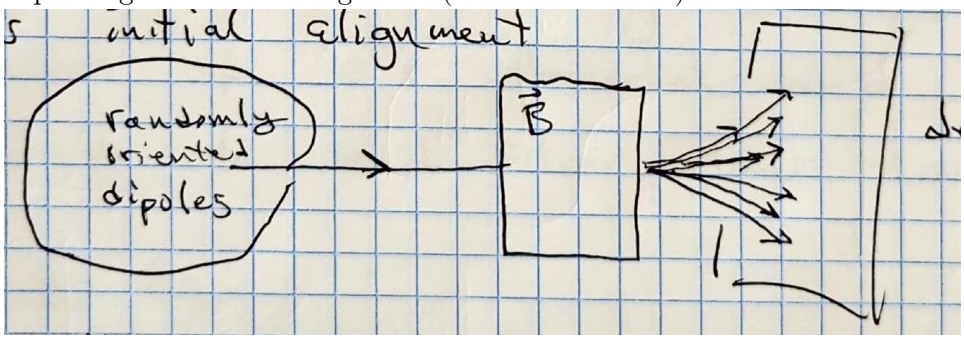
Here we approximate roughly

$$\vec{B} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ zB \end{bmatrix}$$

Classically we know if we sent a little randomly oriented magnet through this vector field then the little magnet should rotate in order to align with the external $\vec{B}$-field viz



Additionally we expect a net force to curve each little dipole up or down depending on its initial alignment (a la Lorentz force)



The path on the highest point of our detector would be all those dipoles in our random sample "beam" which happened to go in perfectly aligned as with the field, We will say "up". The one deflected up the most is called "up".

Notice How crucial it is for $\vec{B}$ to be non-uniform, if the magnetic field were uniform then the forces on the little magnet would cancel out and there would be no deflection.

So now stern-Gerlach ask "Do atoms behave like little magnets or something else?" Other experiments and theories suggested that they do. If an e^- is orbiting a nucleus then that current around the nucleus produces a magnetic field just like Maxwell told us.

Now, SG wanted to study the magnetic properties of e^-'s'themselves and

decided to send silver atoms through $\vec{B}$. Evidently using a cathode ray tube would be much more difficult to set up since the electrons in a beam move fast and deflect very little due to their small masses.

Why Ag?
But Ag has MANY electrons! Why choose Ag if your trying to understand individual electrons?

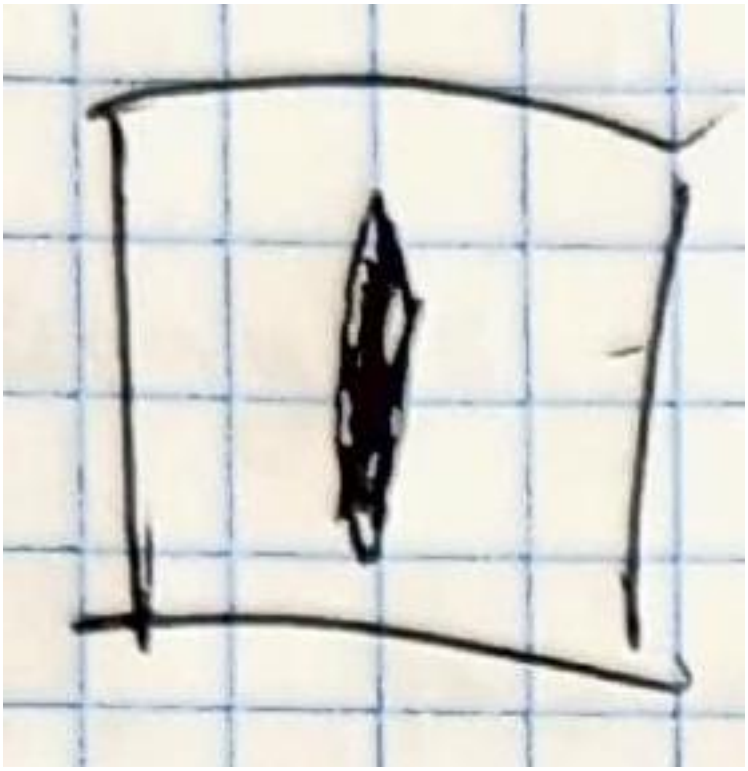
The answer is almost always omitted from textbooks but the reason explains the brilliance of Stern and Gerlach.

Ag was chosen because it is massive and has a single "un-paired" electron with zero angular momentum.

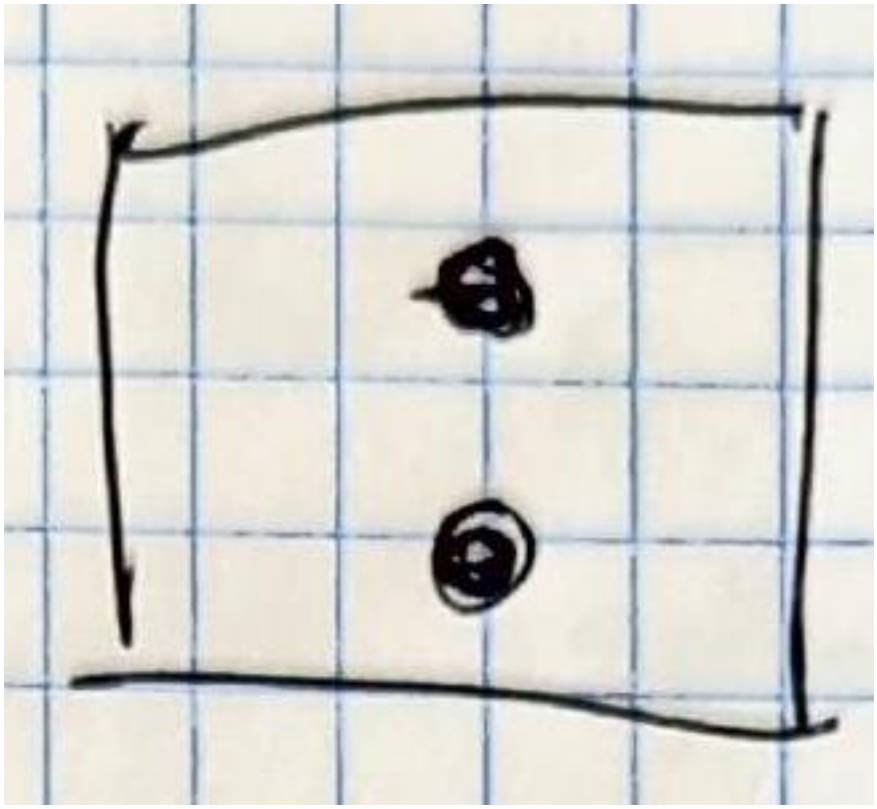
This is the crucial $5s$ electron in silver. Because if all the of the other electrons are paired up then the entire magnetic dipole of Ag should be due to Mr. $5s$! Well, you say, what about protons? Protons indeed but they were too ill understood at the time and had only been discovered about 5 years earlier.

So the silver atoms were absolutely crucial in this experiment!!
So, S&G send in a ton of Ag atoms and use a photographic plate to image how they were deflected after traveling through $\vec{B}$. Everyone expected a continuous vertical line but what was measured looked as though there were just two types of paths! (the actual experiment is blurry for technical reasons but the two paths were clearly obtained).

Expected

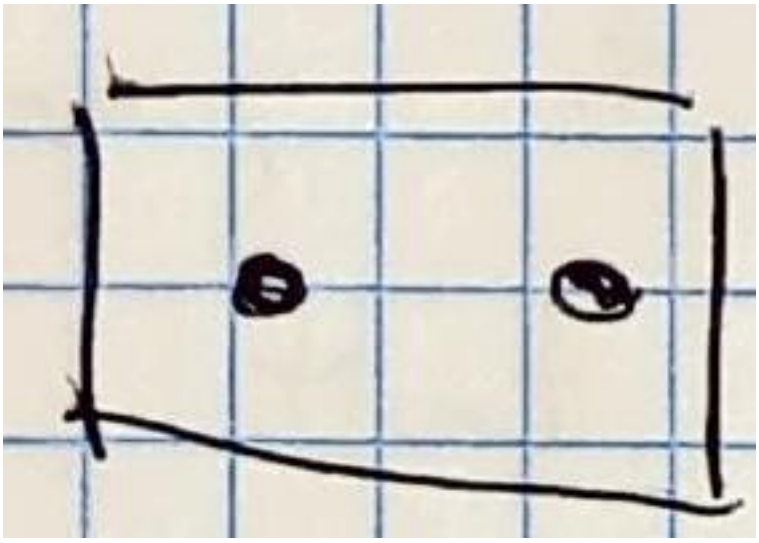


Observed



This is WEIRD! The experiment suggests that the electrons were ALL precisely oriented either up or down when the they went into the experiment. But the A_g atoms were prepared by heating a silver wire and boiling off atoms so this implies there should be a a random collection of atoms/dipoles!

Next they rotated the magnets and observed the same thing: two paths. Evidently nature was conspiring against them if she were ejecting perfectly aligned atoms from their furnace. Now they were perfectly aligned in the other direction!



I can't imagine how profoundly confusing this was at the time. When faced with silly results you wrack your brain trying a bazillion different experiments to verify the silliness isn't the experimenter making a stupid mistake!

No matter what the did they always observed two states. Mystery abound!

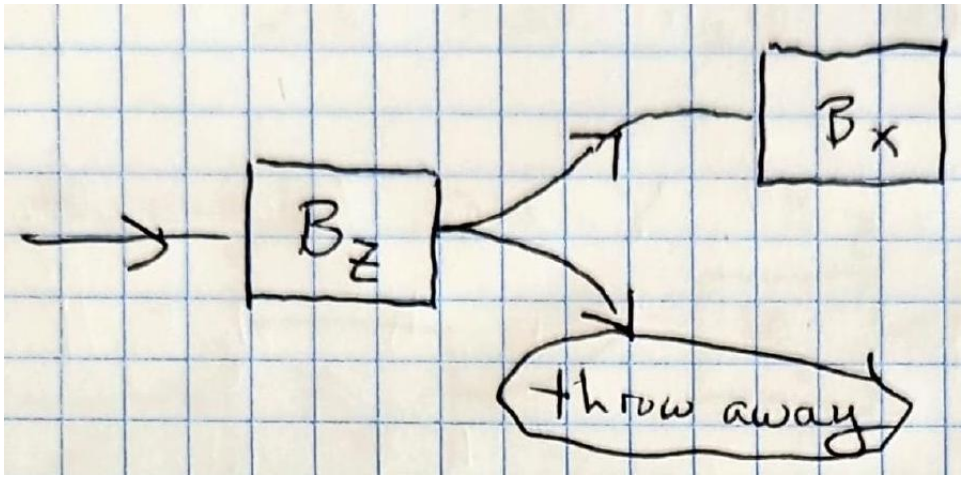
This mystery went unsolved for some four years when Uhlenbeck is Goudsmit offered an interpretation (Konig beat them by a few months but Pauli urged him not to publish such stupid ideas!)

Uhlenbeck is Goudsmit proposed maybe, just maybe, the e^- behaves "as though" it were a spinning ball. This Could produce an "internal" magnetic dipole field.

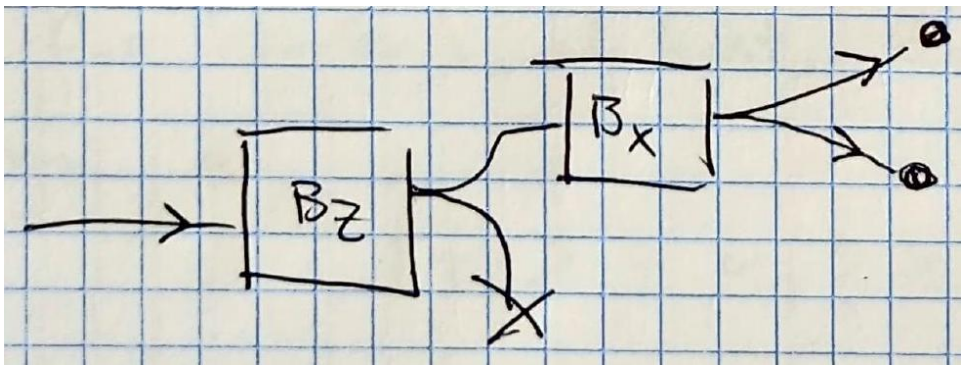
Now, it had been well known for years that if the e^- were a "ball of charge" then it would have to spin faster than the speed of light (Lorentz) in order to maintain it's size and properties. So absolutely NOT a spinning ball of charge!

So clearly U-G are prophesizing blasphemy! Be that as it may it did tentatively explain the experiments if we suppose that the electron has TWO internal states called "up" and "down". But they're weird in that when you measure them they're ALWAYS in either up or down states and never in-between.

e.g. Suppose you take SG experiment and then take all the ups and feed then into ANOTHER SG but with the $\vec{B}$ field pointing
in the x-direction,
viz

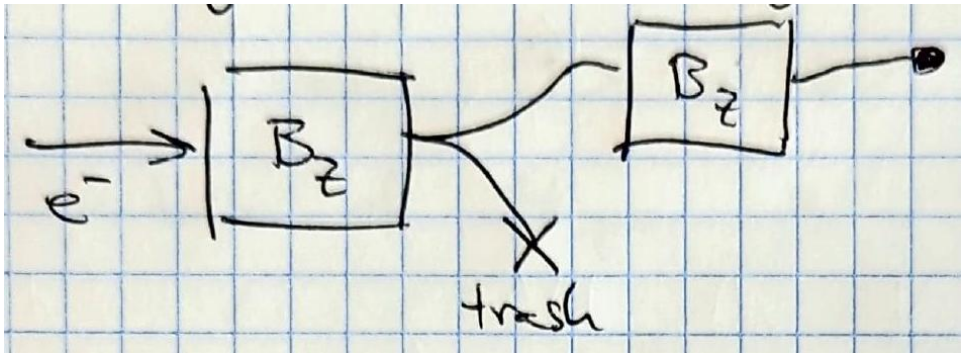


We throw away ALL the down states. There are no more, we got rid of them all. Only up left. We send those up into an x-oriented mung field. So we clearly expect only up coming out the others side.



However, we get a bunch of ups AND a bunch of downs! The ups in state $|\uparrow\rangle$ goes in and then the two states $|\uparrow\rangle, |\downarrow\rangle$ comes out. But how?!

Well then, let's take all the ups and send them through ONLY B_z experiments. so we did a B_z experiment, threw away all the $|\downarrow\rangle$ and did another B_z experiment.



What we find on the other side of the second B_z is ONLY $|\uparrow\rangle$. no $|\downarrow\rangle$'s, zilch.

These "spin states" are very strange.

- Claim: the spin states of an object comprise a Hilbert space

e.g. for electrons, at least, any state can be written as

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$$

In this study we are ONLY interested in spin states. e^-'s have position, momentum, energy, etc. We arent interested in the full Hilbert space; just the "internal" spin Hilbert space. We can tack on the rest of the external variables via $\mathcal{H} = \mathcal{H}_{\text{spin}} \otimes \mathcal{H}_{\text{spacetime}} = \mathcal{H}_{\text{internal}} \otimes \mathcal{H}_{\text{external}}$

Okay. spin states $|\uparrow\rangle$ and$|\downarrow\rangle$ are the basis vectors of $\mathcal{H}_{\text{spin}}$ is a 2 dimensional Hilbert space.

Unitary matrices can change states and observables are required to be Hermitian matrices. Our SG B_z experiment is a measurement operator. We observe two spin states. Theyre observable!

For a 2D space we can write out all possible Hermitian matrices as we have previously shown. viz $M = M^\dagger$

$$M = c_0\mathbb{1} + c_1\sigma_x + c_2\sigma_y + c_3\sigma_z$$

$$\begin{aligned} M &= c_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= c_0\mathbb{1} + c_1\sigma_x + c_2\sigma_y + c_3\sigma_z \end{aligned}$$

These σ_i matrices are called the Pauli Matrices and these are the spin observables.
e.g.

$$\begin{aligned}\sigma_z|\psi\rangle &= \sigma_z(\alpha|\uparrow\rangle + \beta|\downarrow\rangle)\\ &= \alpha|\uparrow\rangle - \beta|\downarrow\rangle\\ \sigma_z\,|\uparrow\rangle &= |\uparrow\rangle\\ \sigma_z\,|\downarrow\rangle &= -|\downarrow\rangle\end{aligned}$$

That is, these $|\uparrow\rangle, |\downarrow\rangle$ states are the eigenvectors of σ_z with eigenvalues ± 1.

We make two algebraic observations

$$\sigma_i^2 = \sigma_i\sigma_i = \mathbb{1} \text{ for all of them}$$

and $[\sigma_x, \sigma_y] = \sigma_x\sigma_y - \sigma_y\sigma_x = 2i\sigma_z$
or generally

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

Where ε_{ijk} is $+1$ if the indices are an even permutation of ijk
-1 if theyre an odd permutation of ijk

This the Levi-Civita symbol.

Notice this looks VERY much like what we found from the generators of a $3D$ vector rotation. There we had

$$[J_i, J_j] = i\varepsilon_{ijk}J_k$$

let us redefine our observables as

$$\hbar\frac{\sigma_x}{2}, \hbar\frac{\sigma_y}{2}, \hbar\frac{\sigma_z}{2}\Big\} = \hat{S}_x, \hat{S}_y, \hat{S}_z$$

Then $\left[\hat{S}_i, \hat{S}_j\right] = i\varepsilon_{ijk}\hat{S}_k$

This is now PRECISELY the Lie-Algebra of vector rotations!

$\Rightarrow \hat{S}_i$ acts like a rotation but in this weird $|\uparrow\rangle, |\downarrow\rangle$ Hilbert space! These aren't your regular rotation matrices.

We say "Spin is an internal or "intrinsic" angular momentum. NOTHING is "Really" spinning here. We are more than welcome to ask the question "are these $\frac{\sigma}{2}i$ the generators of some unitary matrix that is a member of a symmetry group?

Let's see
In classical 3D world we found J_x, J_y, J_z and then exponentiated

$$\mathscr{U} = e^{i\theta\cdot J} = e^{i(\theta_x J_x + \theta_y J_y + \theta_z J_z)}$$

and this performs a vector rotation is $3D$ space. What about this spin business? Well we simply exponentiate the $\hat{S}_i = \frac{\sigma_i}{2}$'s

$$\mathscr{U} = e^{i\frac{\theta\cdot\sigma}{z}} = e^{i\frac{\theta_x}{2}\sigma_x + i\frac{\theta_y}{2}\sigma_y + i\frac{\theta_z}{2}\sigma_z}$$

and this $\mathscr{U}$ acts on these weird $|1\rangle, |\downarrow\rangle$ states. Perhaps this is how $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ changes as we rotate the electron in space. But this $\theta/2$ is funny. What happens to our electron when we rotate by $\theta_z = 2\pi$ we expect nothing. Rotate anything by 2π and watch it go back to itself!

Let's do it;

$$\begin{aligned}\mathscr{U}|\psi\rangle \to \mathscr{U}_{2\pi}|\psi\rangle &= e^{i\frac{2\pi}{2}\sigma_z}|\psi\rangle \\ &= e^{i\pi\sigma_z}|\psi\rangle\end{aligned}$$

Let's check $|\uparrow\rangle \quad \mathscr{U}_{2\pi}|\uparrow\rangle = e^{i\pi\sigma_z}|\uparrow\rangle$

$$= e^{i\pi}|\uparrow\rangle$$

But wait, $e^{i\pi} = -1$ so under 2π rotation $|\uparrow\rangle \to -|\uparrow\rangle$

We don't get $|\uparrow\rangle \to |\uparrow\rangle$ we get up to minus up under full rotation! we have to go $\theta = 4\pi$ to leave $|\uparrow\rangle$ unchanged! We have discovered a new object in mathematics called a "spinor". Some folks think of them as the "square root" of a vector but I personally dont like that intuition.

$$\begin{aligned}\mathscr{U}_{2\pi}|\psi\rangle &= -|\psi\rangle \\ \mathscr{U}_{4\pi}|\psi\rangle &= |\psi\rangle\end{aligned}$$

The states of this Hilbert space are called "spinors". Electrons are described by spinors. These unitary matrices $\mathscr{U}$ are called "generalized rotations"

Spinors exist in our (x, y, z) world. All matter particles in our universe are described as quantum mechanical spinors. Spinors are ultimately a consequence of special relativity, the notion of $3 + 1$ dimensional space-time.

Usually, in physics we represent vectors in $\hat{x}, \hat{y}, \hat{z}$ coordinates and it's so natural we never bothered asking "can I describe these vectors in 1D? 2D, 25D? what would that even mean?!"

It turns out that we CAN! and SHOULD! We can find all sorts of ways by defining a vector as "anything that equals itself after 2π rotation". Or even "a vector is a thing that transforms like a vector". This seems like a circular definition but it's the high brow way of thinking about vectors,

spinors, and especially tensors.
in $SO(3)$ we found

$$[J_i, J_j] = i\varepsilon_{ijk}J_k$$

This expression literally defines what it means to rotate. No mention of what kind of matrices these are: if they behave this way then they are generators of $SO(3)$. The " 3 " in $SO(3)$ naturally causes us to think "these are 3×3 matrices acting on $3D$ vectors. Easy peasy.

These 3×3 matrices and vectors are called the "natural" or "defining" or "vector" representation of $SO(3)$. The defining representations of a symmetry group are generally where we naturally think of starting.

But I can certainly find three 1×1 matrices that satisfy our algebra $[J_i, J_j] = iJ_k\varepsilon_{ijk}$

Try $J_x = 0, J_y = 0, J_z = 0$

Then these J_i satisfy the algebra! B00M I just found a $1D$ representation of $SO(3)$. $SO(3)$ acts on numbers as rotations. That's what this says.

$$\mathscr{U}\, a \in \mathbb{R} = e^{iJ_x\theta_x + J_y\theta_y + J_z\theta_z}a = e^0 a = a$$

The number $13.7 \longrightarrow 13.7$ under rotation by any angle! This is called the "trivial representation". The trivial rep acts on scalars. A scalar is an object that is invariant under rotation! A scalar is an object that transforms the way scalars transform!

For $SO(3)$ we label the possible representations as $j = 0, 1, 2, \ldots$ each representation has dimension of $d = 2j + 1$
e.g. $j = 0$ are the $1D$ scalars. $j = 1$ are the 3D vectors $j = 2$ are the 5D "tensors" and so on. After "vector" literally everything is a tensor. They are just described by ever increasingly many numbers. We wont deal with tensors except maybe $j = 2$ but even then we wont be writing out the actual tensors of that rep!

You may naturally ask "where's the 2D representation of $[J_k, J_j] = i\varepsilon_{ijh}J_k$?". that would correspond to $j = \frac{1}{2}$

Previously we found

$$[S_i, S_j] = i\varepsilon_{ijk}S_k$$

for the $2D$ matrices $\frac{\sigma_x}{2}, \frac{\sigma_y}{2}, \frac{\sigma_z}{2}$

Why didn't they show up in $SO(3)$?! they would be $j = 1/2 \Rightarrow d = 2D$ representations

The reason is that although the ALGEBRAS are the SAME the GROUPS are very importantly DIFFERENT. We've been VERY cavalier by referring to the algebra as $SO(3)$ and often in physics we abuse the mathematics that way. Mathemeticians have distinguishing notation for LA and Groups.

They says the GROUP is $SO(3)$ and the Lie Algebra is $\mathfrak{so}3$. This is the Fraktur font. the Lie Algebra is technically the "Tangent Space" of the Lie Group near the identity $\mathbb{1}$. My god so many fonts! we will take care to distinguish the LG from the LA but I will certainly forget to distinguish here in there as I'm a low-life physicist instead of a high-minded mathematician.

Let's discover what's so different about $SO(3)$ and $SU(2)$.

In the defining representation of $SO(3)$ we have the mighty 3-component vector that rotates around and around under our rotation matrix ($R_{\vec{\theta}}$ is an element of the $SO(3)$ group).

$$\vec{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \quad \vec{v} \to \vec{v} \text{ under } 2\pi \text{ rotation}$$

CLAIM: I can package v_x, v_y, v_z into an $\mathfrak{su}2$ matrix as follows:

$$\begin{aligned} M \in \mathfrak{su}2 \to \vec{v}\cdot\vec{\sigma} &= v_x\sigma_x + v_y\sigma_y + v_z\sigma_z \\ &= \begin{bmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{bmatrix} \end{aligned}$$

this is a 2 by 2 matrix inside the Lie Algebra $\mathfrak{su}2$
Sometime this package is confusingly called a "Pauli vector".

Now, matrices and vectors "transform differently under rotations". They're different objects after-all!

For $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ we apply a matrix $R(\vec{\theta}) = R$ (we will suppress the angles) to get $\vec{y}$

$$\vec{y} = R\vec{x}$$

"States" or "vectors" transform as

$$\left|\psi'\right\rangle = \mathscr{U}\left|\psi\right\rangle \tag{1}$$

but Matrices/Operators transform as

$$\tilde{A} \to \mathscr{U}^\dagger A \mathscr{U} = \mathscr{U}^{-1} A \mathscr{U} \tag{2}$$

Matrices get sandwiched between $\mathscr{U}$'s'. If we take this Pauli package and we wish to rotate the matrix package then we have to follow rule (2). Let $\vec{\alpha}$ represent the rotation angles.

$$\mathscr{U} = e^{i\frac{\vec{\alpha}}{2}\cdot\vec{\sigma}}$$

$$\mathscr{U}^\dagger = \mathscr{U}^{-1} = e^{-i\frac{\vec{\alpha}}{2}\vec{\sigma}}$$

$$\Rightarrow \mathscr{U}^\dagger \begin{bmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{bmatrix} \mathscr{U}$$

$$e^{-i\frac{\vec{\alpha}}{2}\vec{\sigma}} \begin{bmatrix} v_z & v_x - iv_y \\ v_x + iv_y & -v_z \end{bmatrix} e^{i\frac{\vec{\alpha}}{2}\cdot\vec{\sigma}}$$

This $\mathscr{U}^\dagger A \mathscr{U}$ does the same operation as our regular 3×3 Rotation matrix acting on a vector

$$\hat{R}(\alpha)\vec{v} = \mathscr{U}^\dagger_{\frac{\alpha}{2}}\vec{v}\cdot\vec{\sigma}\mathscr{U}_{\frac{\alpha}{2}}$$

The half-angles in $SU(2)$ combine into a FULL angle in $SO(3)$ and everything watches up!

Now, investigating further. For simplicity consider unit vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{x}$

i.e. $x^2 + y^2 + z^2 = 1$

Under $SO(3)$ and also under $SU(2)$

$$X = \vec{x}\cdot\vec{\sigma}$$

X is our Pauli vector or package but now it's a unit vector package.

Now, for $SO(3)$ we required

$$\begin{cases} \det R = 1 & \text{(special)} \\ R = R^T & \text{(orthogonal)} \\ J = -J^T & \text{(antisymmetric generators)} \end{cases}$$

for $SU(2)$

$$\begin{cases} \det \mathscr{U} = 1 & \text{(special)} \\ \mathscr{U}^\dagger \mathscr{U} = \mathbb{1} & \text{(unitary)} \\ S = S^\dagger & \text{(Hermitian generators)} \end{cases}$$

Now, you can show that

$$\det X = -\vec{x}^2 = -\vec{x} \cdot \vec{x} = -\left(x^2 + y^2 + z^2\right)$$

under $SU(2)$ rotation

$$X' = \mathscr{U}^\dagger X \mathscr{U}$$

If X is Hermitian then X' is Hermitian

Proof:

$$X'^\dagger = \left(\mathscr{U}^\dagger X \mathscr{U}\right)^\dagger = \mathscr{U}^\dagger X^\dagger \mathscr{U} = X'$$

$$\therefore X'^\dagger = X'$$

Now, we ask for

$$\det X' = \det\left(\mathscr{U}^\dagger X \mathscr{U}\right) = \det \mathscr{U}^\dagger \det X \det \mathscr{U}$$

but $\det \mathscr{U} = 1$

$\therefore \det X' = \det X$

or $\vec{x}'^2 = \vec{x}^2$

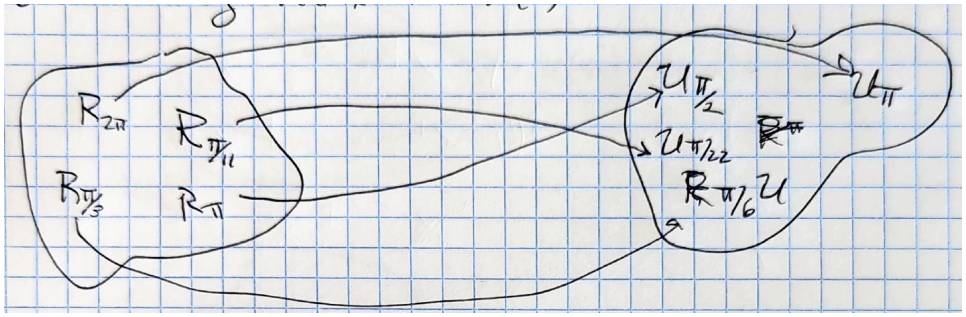
i.e. lengths dont change under $SU(2)$ just like for $SO(3)$ rotations!

viz, $\vec{x}' = R\vec{x} \Rightarrow \vec{x}^2 = \vec{x} R^T R \vec{x} = \vec{x}^2$

So we discover that we can perform rotations with $SO(3)$ or $SU(2)$. It walks and talks like a duck under both groups!

Question: For every rotation matrix do we have a unitary matrix in in $SU(2)$?

Answer: YES! but...



Firstly: There does exist a function $f : \mathrm{SO}(3) \longrightarrow SU(2)$. This takes an $SO(3)$ rotation and spits out an $SU(2)$ matrix that does the same job of rotating a vector.

However this is NOT bijective! If it were a bijective function then what would be the point. They'd be isomorphic and for all practical purpose no difference whatsoever. $SO(3)$ and $SU(2)$ however, are VERY different in subtle ways. Notice

$$\vec{x}' = R\vec{x}$$

$$X' = \mathscr{U}^\dagger X \mathscr{U}$$

$\vec{x}$ and X are EXACTLY the same information just packaged in different ways. but notice that

$$X' = (-\mathscr{U})^\dagger X(-\mathscr{U}) = X' = \mathscr{U}^\dagger X \mathscr{U}$$

There exist TWO $SU(2)$ matrices for every $SO(3)$ matrix!

We say $SU(2)$ "double covers" $SO(3)$ but they are locally isomorphic (i.e. their Lie Algebras are isomorphic). viz

$$\mathfrak{so}3 \xrightarrow{isomorphic} \mathfrak{su}2$$

Because, just look at their Lie Algebras,

$$[J_i, J_j] = i\varepsilon_{ijk}J_k$$

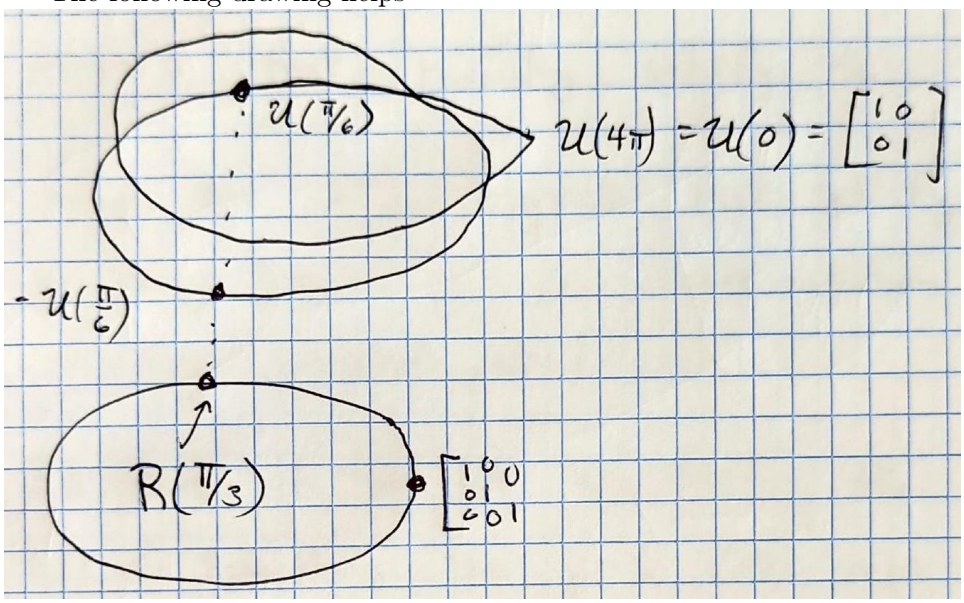
$$[S_i, S_j] = i\varepsilon_{ijk}S_k$$

They're identical. The algebras are the same. The groups certainly are NOT the same

They are locally isomorphic (the Lie Algebra is the tangent space of the group near $\mathbb{1}$)

but the groups are globally different!

The following drawing helps



As we walk all the way around in $SU(2)$ we go around twice in $SO(3)$. Spinors have to rotate in our space by 4π to go back to where they started. Right, this is bizarre. I don't "see" anything behaving this weirdly!

Now, if Hilbert space is required to have unitary states the objects in the theory must be unitary representations of our beloved rotational symmetry.

Objects that obey rotational symmetry are eigenstates of the $\mathfrak{su}2$ generators

$$\left[\frac{\sigma_x}{2}, \frac{\sigma_y}{2}\right] = \frac{i\sigma_z}{2} \Rightarrow \left[\hat{S}_i, \hat{S}_j\right] = i\varepsilon_{ijk}\hat{S}_k$$

We can easily find a 1D set of objects the scalars that trivially satisfy this

We also found the 2D objects that satisfy this and they are the $j = 1/2$ representation (e.g. electrons, neutrons, etc) of $\mathfrak{su}2$

The 3D reps of the $\mathfrak{su}2$ LA are NOT the (x, y, z) of something. The $j = 1$ (3D) reps of the $\mathfrak{su}2$ LA are quantum Objects/states. They are "internal". They dont manifest observationally perse, but they have massive consequences.

for $j = 1$ the states are

$$|-1\rangle, |0\rangle, |1\rangle$$

We call these spin-1 states (this is almost a photon).
Another example of spin-1 is a "cooper pair" which are responsible for superconductivity.

$j = 3/2$ states are

$$|-3/2\rangle, |-1/2\rangle, |1/2\rangle, |3/2\rangle$$

the Ω-particle is spin $3/2$. Some Atoms and Nuclei fall into these representations.

$j = 2$ has to do with gravity.

In general the eigenstates of Spin for the j-representation are

$$|-j\rangle, |-j+1\rangle, |-j+2\rangle, \ldots ...|+j\rangle$$

Most of physics is apropos spin-$1/2$ and spin-1. The half-integer objects are call Fermions and the integer objects are called Bosons and they have strikingly different behavior.

Yet again let us remind ourselves of the journey we're on. We've covered a whirlwind of complicated abstract mathematics an physics and it seems we are no closer to actually solving a damn problem we might face in the real world. We shall get there in due course and based on my personal relationship and struggles with quantum theory I have brought you along this particular path towards reality simply because I hope that it will save you the multi-years long struggle and stress I, myself, had to overcome before some of these things began to "make sense", whatever that means.

Recap

1. We postulated reality is based on complex vectors in a Hilbert Space. We had ZERO reason to do this other than hindsight.
2. We can't predict WHAT exactly will happen in QT but we CAN find probabilities for any possible experimental result.
3. We expect the universe to have some basic symmetries we all tend to take for granted.

4. Following Lie's advice we generated those symmetries by operators/matrices J, L, etc

5. The eigenstates of the generators are particular states that we expect to be invariant and the eigenvalues are the measurable quantities of those states.

6. We claimed that each "Group" of symmetries can be "represented" as some vector space. From this we found Lie Algebras and claimed to find MANY MANY MANY possible representations of the Group. Sometimes we wish to rotate vector quantities such as (x, y, z) and other times we wish to rotate more complicated objects called tensors and these require larger vector spaces to represent rotations.

7. We finally found ALL representations of $SO(3)$ and of $SU(2)$ where $SU(2)$ described the unusual "internal" states of some quantum objects called fermions.

8. We then found that there's a deep relationship between $SO(3)$ and $SU(2)$ and even found a local map between the two groups. In the sense we've seen $SU(2)$ subsumes everything about $SO(3)$. We found not only $\ell = 0, 1, 2, \cdots$ but also half-integer representations of rotations and our universe makes use of those $s = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$ representations. Evidently we discover the symmetry of our universe is not "just" $SO(3)$ but rather more generallly $SU(2)$ as Stren, Gerlach, Goudsmit, Pauli, Koning, and Uhlenbeck discovered so remarkably!

9. The groups of translations, generalized rotations are some of the primary symmetries of our universe. There are others such as $SO(3, 1)$ - called the Lorentz Group and there also discrete symmetries we expect that we haven't covered such as "Parity - P", "time reversal - T", and "charge conjugation - C". As far as we currently understand the symmetries of the universe are the Poincare Group and CPT.

Our next task is to finally begin doing some firm physics. I will teach you how to write our generator of time translations $\hat{H}$ in a way that connects with what you have previously learned classically. We will use this specific $\hat{H}$ to solve MANY problems in quantum theory. Almost everything revolves around using $\hat{H}$ to write and then solve a particular differential equation either algebraically or by using the tools of calculus. We shall see examples of both!

Before that, I think we should take a step back into classical land and discuss the various methods we have to solve classical systems. By now you are extremely familiar with Newton's second law but this methodology is rather clunky and difficult for systems much more complex than a series of pulleys and blocks. There are three other formalisms entirely parallel to Newton's method that allow us to solve much more complicated and much more general systems which completely bypass the notions of force and instead focuses on energy. We shall study the following methods rather briefly in order to motivate our quantum formalism.

1. Hamilton' Mechanics
2. Lagrange's Mechanics
3. Action

Hamilton's Mechanics

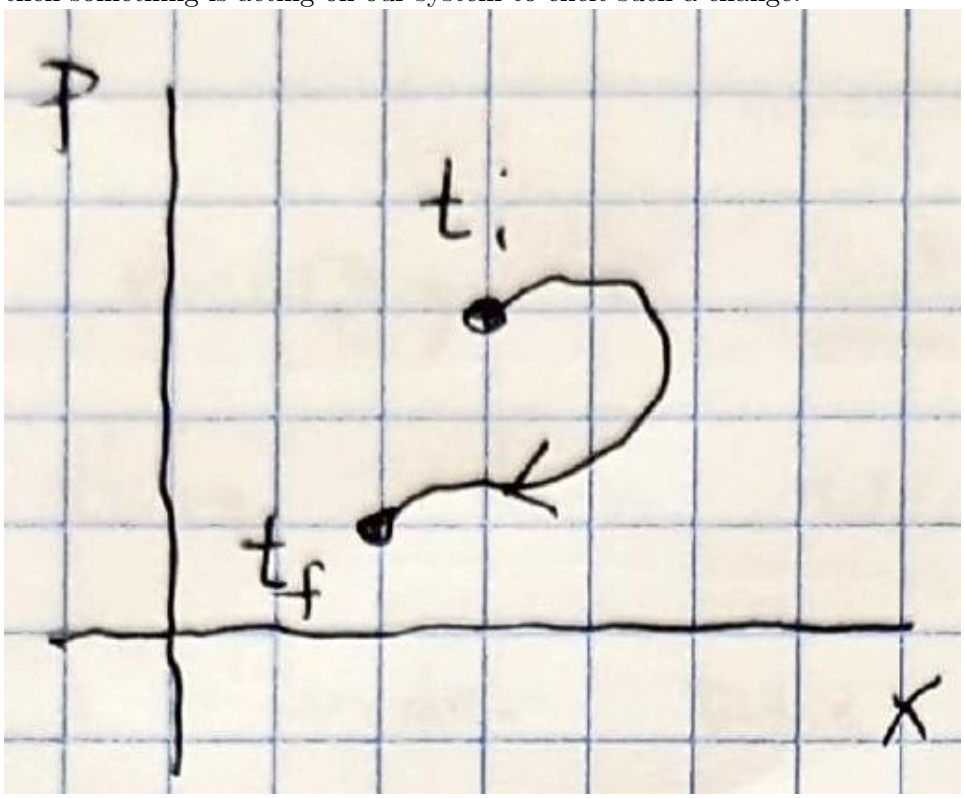
What the heck is a state any way?

Laplace pointed out that a consequence of Newton is that if you have a system and somehow know at some point in time every position and every velocity(to supreme accuracy) then it follows that you can predict (to supreme accuracy) the positions and velocities at all future times. Laplace took this even further and found that classical physics implies that the future of the entire universe can be accurately predicted from initial data. This is called Laplace's demon. Around the turn of 1900 Poincare realized this might not be the full story and in the 1960s Lorenz was building simple Newtonian-like models of weather patterns such that trajectories of infinitesimally different initial conditions could differ wildly. Thus was born the subject of Chaos Theory. But, still Laplace's demon holds if the demon possesses perfect accuracy - meaning that inside the demon's "brain" can be stored $\mathbb{R}$ number quantities (or at least algorithms for their construction!): something that modern computers can never hope to do! Everything a physicist measures is a rational number in $\mathbb{Q}$.

Anyhow, because of this dynamical determinism we define a classical "state" to be the list of all positions and all momenta at some point in time. It is represented by a single point in the space of all positions and all momenta. For a fly buzzing around the room this point is in a 6-dimensional "space". For a bunch of buzzing flies, say N, this "phase space" is $6N$

dimensional. A room full of gas would be 6×10^{23} dimensional! But that's okay: a single point is simply a single point. You merely place a coordinate system with axes representing positions and velocities/momenta (see below).

This single point in phase space represents what we call a single classical "state". At some later time the "state" may have changed somehow according to some function(s) that determine the rule for such a state change. There could be a continuous change as a function of time or even some other parameter (say, magnetic field). At any rate, if we find our state has changed then something is acting on our system to elicit such a change.



Now, I think we all agree that time is very special and if t is "more special" than any other parameter λ then there should be, we hope, some special function of (p, x) that takes an input state and outputs some other state. We may even suspect this function is a continuous mapping and states evolve along some special "time trajectory". The paths that occur in phase space are called trajectories.

We will not derive this special function but only point out what we previously learned.

In quantum mechanics, recall, we found that, if the energy is not changing

as a function of time then we expect the system to be "invariant" under time translation.

From this germ of an idea we found the "generator" of time-translation to be given by an operator $\hat{H}$ such that
and

$$|\psi(t)\rangle = e^{\frac{i\hat{H}t}{\hbar}} |\psi(t=0)\rangle$$
$$\hat{H}\,|E_i\rangle = E_i\,|E_i\rangle$$

implying that the eigenvalues of H are the energies of the quantum object.

OKAY, classically we know how to write down the total energy of a system.

$$E = \frac{p^2}{2m} + V(x) = KE + PE$$

For a constant energy value we anticipate some change in the state. e.g. if there's no potential but we have a kinetic energy then our object will clearly be changing its x position along the x-axis at constant momentum. This would be a horizontal line and it would evolve in time. That sounds like the special function we are looking for that generates a state transformation in time! This is the special function we sought that takes an initial condition to a final condition along a specific trajectory. We will call this energy $E = H$ where for E we mean a particular value, say $E = 5J$ but H is a function of p and x and can have any value (just like if we had a math problem where $z = f(x, y)$ that represented the height above the xy plane for some function). We postulate, with reasonable motivation, that

$$H(p, x) = \frac{p^2}{2m} + V(x) = KE + PE$$

Technically the potential could be also a function of $p, V(x, p)$ or $V(p)$ but this rarely happens (ever?) in practice.

BOLD CLAIM

If you know $H(p, x)$ (called "the Hamiltonian")and the initial state (p_i, x_i) then you know everything about the evolution of your system!

Hamilton told us how to solve a system using this phase space and total energy function. We simply solve a pair of differential equations using the partial derivatives of the Hamiltonian:

$$\dot{\vec{x}_i} = \frac{\partial H}{\partial \vec{p}_i}$$

$$\dot{\vec{p}}_i = -\frac{\partial H}{\partial \vec{x}_i}$$

Note: here i represents each particle of our system and this notation means that each component of the vector is given by, for example, the y component for a single mass

$$\dot{y} = \frac{\partial H}{\partial p_y}$$

So these two equations are actually $3N$-pairs of equations where N is the number of objects in your total system (e.g. a gas of molecules or something)

This is entirely equivalent to Newton's Laws but it is much more general. This method is called "Hamiltonian Mechanics".

What's more, this route leads directly to the statistical mechanical foundation and explanation of the whole field of thermodynamics following Boltzmann's work.

We shall next build our intuition of this phase space by trying to figure out what the trajectories should look like in phase space for a familiar problem. We will NOT solve this problem just yet: we only want to sketch what we think the solutions should look like in this new abstract space. Then we will solve the system using Hamilton's Equations and find exactly what we find using Newton!

Example: A pendulum, (we shall write x as the angular/position variable)

A single pendulum has

$$KE = \frac{p^2}{2m}$$

and

$$PE = -mg\cos x = -A\cos x \text{ with } A > 0$$

$$\therefore \quad H(p,x) = \frac{p^2}{2m} - A\cos x$$

What do the trajectories in phase space look like for some particular value of $E = H(x,p)$?

Suppose our pendulum has total energy $E < A$ then (converting to $\dot{x}, x$ for convenience)

$$E = \frac{1}{2}m\dot{x}^2 - A\cos x$$

Physically if E is small then we expect that the pendulum will be oscillating only slightly back and forth (a small angle x) constantly transferring its energy between kinetic and potential.

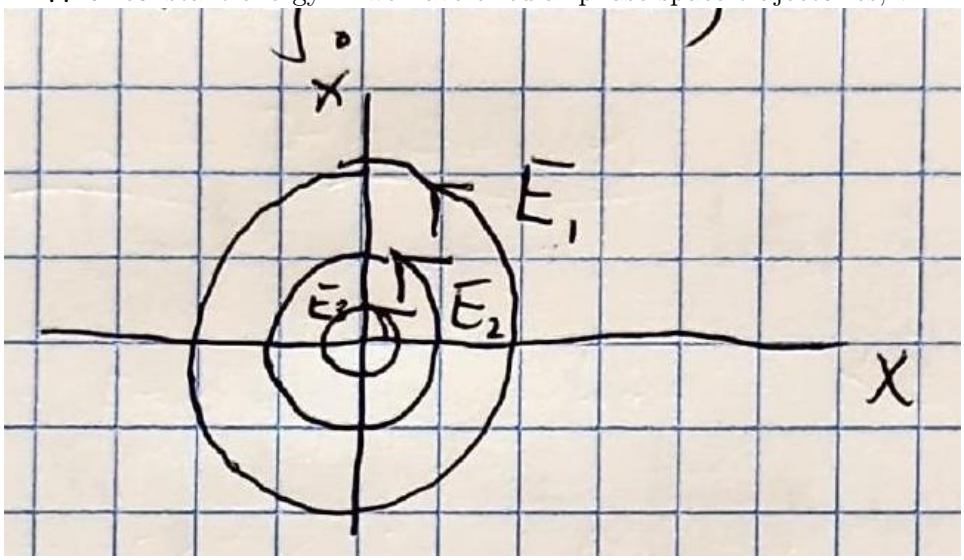
i.e. $x \ll 1$.

If this is the case then we can Taylor expand the potential $\cos x \approx 1 - \frac{x^2}{2!} + \ldots$ and keep x^2 term
then

$$E \approx \frac{1}{2}mx^2 - 1 + \frac{x^2}{2}$$
$$E + 1 \approx \frac{1}{2}mx^2 + \frac{1}{2}x^2$$

In phase space for $m = 1$ we can plot $(\dot{x}, x)$ coordinates then this $E + 1$ represents a circle of radius $\sqrt{E+1}$

$\therefore$ for constant energy E we have circular phase space trajectories, viz



If we had a single perfect circle for each E we would say" E foliates the phase-space with circles". However, these are only approximately circles and only valid for small angles x. At larger values of x they become distorted and may even turn into other trajectories.

Anyhow, our $H(p, x)$ describes changes in (p, x) but those changes are such that we always stay on ONE constant Energy circle. The pendulum lazily swings back and forth. Pick ANY point on this constant energy surface and start your experiment then you will watch the initial point trace out the circle corresponding to that energy. As the angle x changes the velocity will

change accordingly. when the angle is 0 then we expect the SPEED will be maximum. The velocity will point either to the left or the right depending which way the mass is moving. But thats precisely the behavior of these circles!

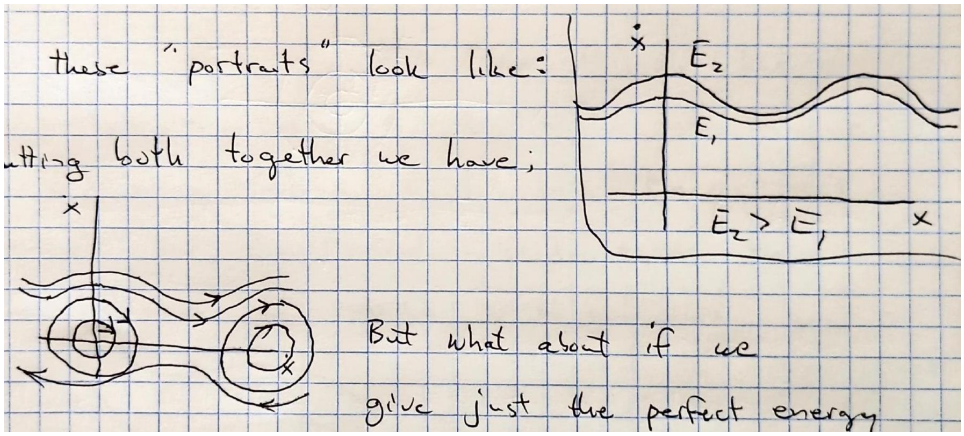
Now, clearly this isn't the entire phase space. we can also have $E = A$ and $E > A$

Physically we can input lots of energy into our pendulum and cause it to wrap around its origin/pivot forever. The speed is maximum at the very bottom of the pendulum and minimum at the very top. With a perfect amount of energy we could cause the pendulum to swing towards the top while slowing down such that it comes to rest exactly at the top (this is a rod-like pendulum rather than a string). Unstable indeed! Any slight disturbance from the top will cause it to fall into a circle-like trajectory.

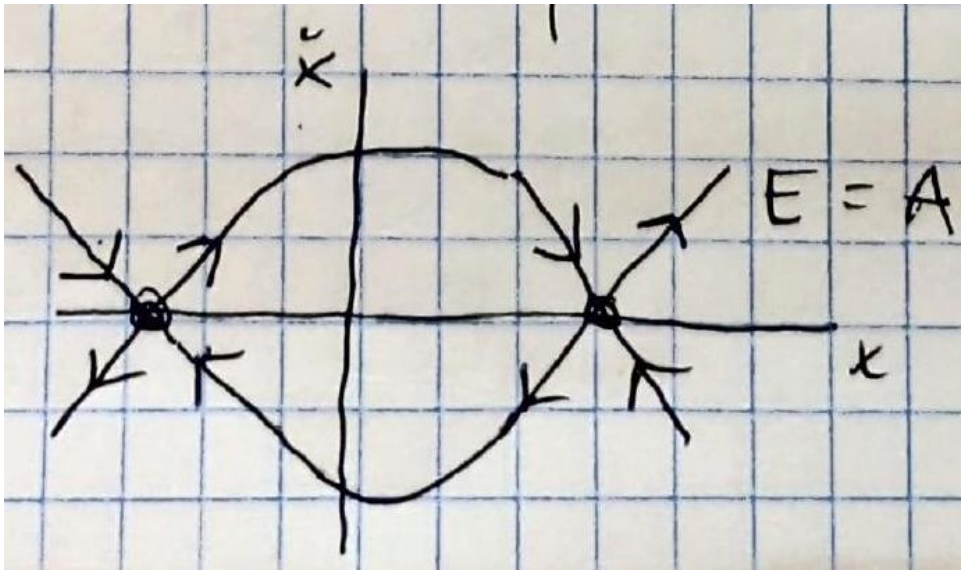
Since x measures the angle of the pendulum it varies from 0 radians to 2π radians in one rotation. But with large energy we can make it rotate as many times around as we like.

Therefore, the range of x is technically infinite. Also, note, because there is no dissipative forces the energy is always constant so if we just barely get over the top it will keep rotating around and around forever.

So far, we expect these "portraits" of trajectories to look like:

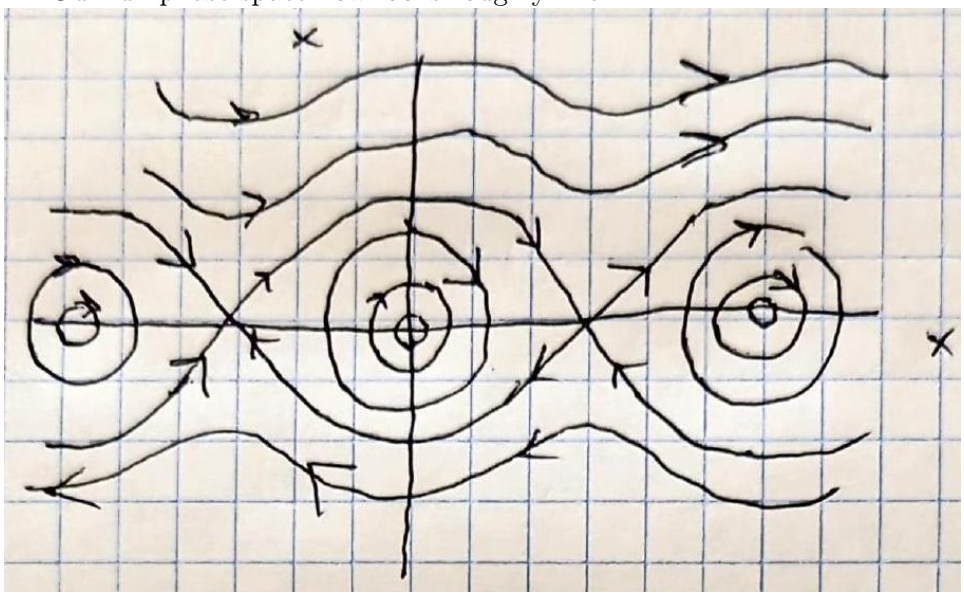


And if we give just the perfect precise amount of energy such that the pendulum gets to the top and stays there we will have trajectories that cross at points represented by the top position x_{top}. This is not quite realistically possible but mathematically we should have the possibility otherwise our phase space wouldn't be "foliated" (there would be missing trajectories). These trajectories look like



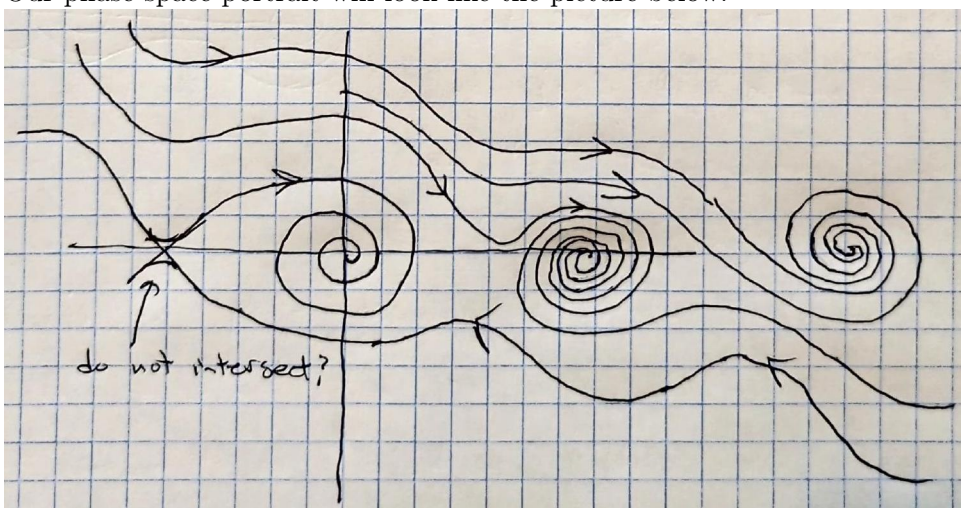
Pay attention to the the arrows, they tell us how the point in phase space is moving as a function of time. These crossings are unstable Equilibria

Our full phase space now looks roughly like:



The trajectories foliate the entire space. There is a single trajectory for every value of E. Every single point in phase space corresponds to one and only one trajectory. This is not a priori obvious mathematically but seems reasonable physically. Afterall H was built out of x and p.

We could even imagine friction being involved. In this case Energy isn't constant and the pendulum would eventually settle down to some minimum with zero velocity. It would spiral in to special points which represent the bottoms of the pendulum! EVERY energy trajectory will terminate at some bottom position whether it swings around and around 100 times or just one. Our phase space portrait will look like the picture below.



Your task is to find the phase space portrait for the Earth orbiting the Sun.

Onward to solving the simple harmonic oscillator (not the pendulum).

We wish to solve

$$\dot{x} = \frac{\partial H}{\partial p}$$

$$\dot{p} = -\frac{\partial H}{\partial x}$$

For the Hamiltonian given by

$$H(p, x) = \frac{p^2}{2m} - A\cos x$$

$$\dot{x} = \frac{\partial}{\partial p}(\frac{p^2}{2m} - A\cos x)$$

$$= \frac{p}{m}$$

and

$$\dot{p} = -\frac{\partial H}{\partial x} = -\frac{\partial}{\partial x}(\frac{p^2}{2m} - A\cos x)$$

$$= -A\sin x$$

but

$$\frac{d^2x}{dt^2} = \frac{d}{dt}\frac{p}{m} = \frac{\dot{p}}{m} = -\frac{A\sin x}{m}$$

or

$$m\ddot{x} = -A\sin x$$

which is what we would expect from Newton's second law. To solve this non-linear differential equation we would probably want to make a small angle approximation! Otherwise it has no known analytic solutions! Numerical solutions only!

Let us solve the most important system in all of physics: The Simple Harmonic Oscillator (SHO).

For a spring system we have $KE + PE$ for some stretch from equilibrium as,

$$H(p,x) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2$$

Hamilton says "take the respective partial derivatives of H".

$$\dot{x} = \frac{\partial}{\partial p}(\frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2)$$

$$= \frac{p}{m}$$

and

$$\dot{p} = -\frac{\partial H}{\partial x} = -\frac{\partial}{\partial x}(\frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2)$$

$$= -m\omega^2x$$

but,

$$\frac{d^2x}{dt^2} = \frac{d}{dt}\frac{p}{m} = \frac{\dot{p}}{m} = -\omega^2x$$

From this we find

$$\ddot{x} = -\omega^2 x \Rightarrow \ddot{x} = -\frac{k}{m}x$$
$$\Rightarrow m\ddot{x} = -kx$$

Which, again, is Newton's 2^{nd} Law! So why bother with Hamilton?

The whole point of the Hamilton formulation is to trade a single 2^{nd} order differential equation with two 1^{st} order coupled differential equations.

1^{st} order equations are typically MUCH easier to to solve even if they are coupled equations. Furthermore, complex systems with many interacting parts are generally easy to solve with Hamilton and near impossible with Newton.

The solutions to $\ddot{x} + \omega^2 x = 0$ directly from Hamilton's Equations are

$$x(t) = Ae^{\pm i\omega t}$$
$$p(t) = m\dot{x}$$

just as we expect (okay, so you have to take the Real part of these complex valued equations for physical solutions)!

There's MUCH more to say classically about the oscillator and Hamilton but we will now switch to Lagrange and Action.

Action is where the physics is at

Functionals

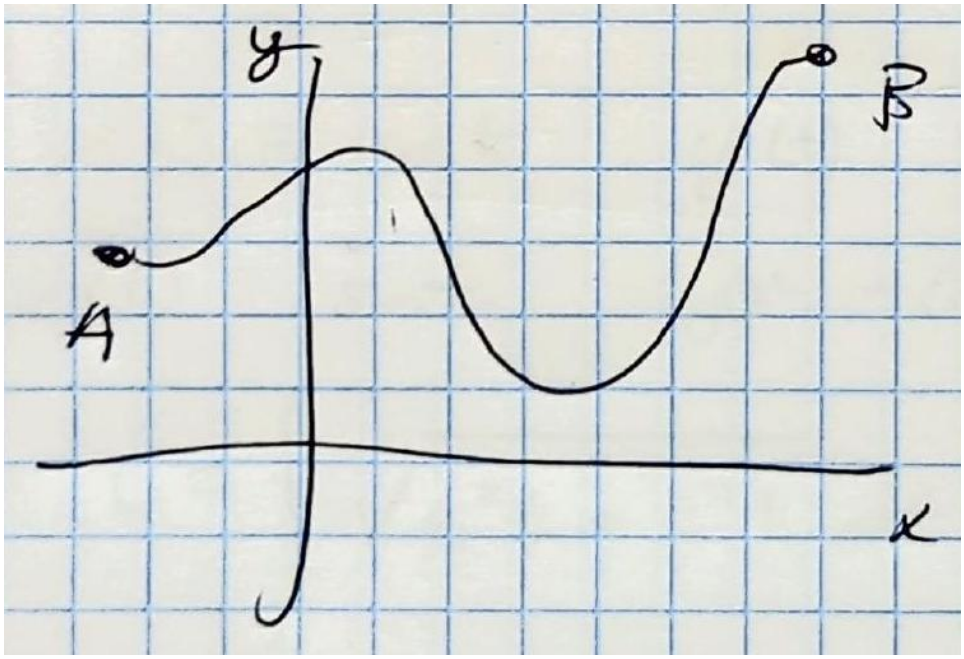
A functional is a mathematical object that maps functions to real numbers. it is denoted with square brackets as follows:

$$\mathscr{L}[f] \in \mathbb{R} \quad \mathscr{L} : f \to \mathbb{R}$$

How might this happen? What are some examples?

Example 1: The total length of a curve.

Suppose we have a curve in the x-y plane as below



This curve can be parameterized however we wish, usually with length l or time t. This just tells us where we are on the curve.

We can find the total length by adding up all the incremental lengths ds given by Pythagoras

$$ds^2 = dx^2 + dy^2$$

Then total length is just the integral of ds over the parameter space

$$L = \int ds = \int_a^b \sqrt{dx^2 + dy^2}$$

Now, if we have a path given by some specific parameterization $x(t), y(t)$ then when we input these curves into out total length integral out pops a $\mathbb{R}$ number length. That sounds like L is a functional!

For example,

$$x(t) = \cos t \quad y(t) = \sin t$$

then

$$\dot{x}(t) = \sin t \quad \dot{y}(t) = -\cos t$$

then

$$\dot{x}^2 + \dot{y}^2 = 1$$

$$L = \int \sqrt{\dot{x}^2 + \dot{y}^2} dt = \int_{t_A}^{t_B} dt = t_B - t_A$$

This says the total length is a functional of $x(t), y(t)$
i.e.

$$L[x(t), y(t)] \in \mathbb{R}$$

Different curves will yield different lengths. This is more obvious than obvious. duh!

Another functional used extensively in physics is called the Action. Don't ask me what Action is. I have no idea! We do know Nature LOVES action. Mysterious! We say,

$$S : q(t), \dot{q}(t) \to \mathbb{R}$$

where $q(t), \dot{q}(t)$ is a trajectory in phase space and S is given by

$$S[q(t), \dot{q}(t), t] = \int \mathscr{L} dt$$

where $\mathscr{L} = KE - PE$ is called the "Lagrangian" which is the difference between kinetic energy and potential energy of a physical system. For example, typically,

$$KE = \tfrac{1}{2} m\dot{q}^2 \quad PE = V(q)$$
$$S[q, \dot{q}] = \int dt \left(\tfrac{1}{2} m\dot{q}^2 - V(q) \right)$$

yields different values for different trajectories $q(t), \dot{q}(t)$. Additionally, these trajectories could very well be multi-dimensional (i.e. 3 dimensional space paths).

Example:

Kinematics - constant acceleration and no potential energy (a free particle).

From Galileo we recall,

$$
\begin{aligned}
x(t) &= x_0 + v_i t + \frac{1}{2} a t^2 \\
\dot{x}(t) &= v_i + at \\
S[x, \dot{x}] &= \int dt \frac{1}{2} m \left(v_i + at\right)^2 \\
&= \frac{1}{2} m \int_p^2 v_i^2 + 2 v_i a t + a^2 t^2 \\
&= \frac{1}{2} m \Bigg(v_i^2 \left(t_b - t_{da}\right) + v_i a \left(t_b - t_a\right)^2 + \frac{a^2}{3} \left(t_b - t_a\right)^3
\end{aligned}
$$

take $m = 1, v_i = 0, a = 1$

$$
S = \frac{1}{6} \left(t - t_a\right)^3
$$

The physical units of action is Js (Joules seconds). Again, i have no idea what action is or why nature utilizes it but for Galileo this value is the action. What we DO with the action makes ALL the difference. It's MAGICAL!!

But first, let's go back to curve length and ask an obvious question: "What is the specific curve that minimizes the distance between two given points in a given flat Euclidean space?"

There exist MANY possible curves between two given points. In flat space we expect that the minimum length curve between points A and B is a straight line $y = mx + b$. Can we prove this using a functional?

Idea: Let us find a way to "extremize" $L = \int dt \sqrt{\dot{x}^2 + \dot{y}^2}$. If L is minimal then there will be a specific $x(t)$ and $y(t)$ associated with that length (perhaps even an entire family of curves all minimizing L! This is very similar to finding the maxima and minima in Calculus class where you set the first derivative to zero and then check the second derivative.

Let's rewrite our general curve as $y(x)$ and then integrate over x to find the total length

$$
L = \int_{x_A}^{x_B} dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \int dx \sqrt{1 + f'(x)^2}
$$

Our functional is of the form

$$
L = \int_a^b dx F\left(y, y', x\right)
$$

Now, we ask "if we change the path $y(x)$ by just a little bit (called $\delta y(x)$ - a small arbitrary change at each point x), then how will L respond and change?"

We will consider variations such that the end points A and B will remain fixed. i.e. $\delta y(x = A) = \delta y(x = B) = 0$.

$$L \longrightarrow \delta L = \int_A^B dx \left[F\left(y + \delta y, y' + \delta y'\right)\right]$$

now $F\left(y + \delta y, y' + \delta y'\right)$ looks like it should be expanded via Taylor to first order if these variations are small which indeed they are ($\delta y << 1$)

$$\delta F = \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y'$$
$$\therefore \quad \delta L = \int_a^b dx \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y'$$

But now what? we have δy and $\delta y'$ and

$$\delta y' = \delta \frac{dy}{dx}$$

I wont prove this by it can be shown (do it!) that derivative and variations commute

$$\delta \frac{dy}{dx} = \frac{d}{dx}\delta y$$
$$\therefore \delta L = \int dx \left(\frac{\partial F}{\partial y} + \frac{\partial F}{\partial y'}\frac{d}{dx}\right)\delta y$$
$$\delta L = \int dx \left(\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\frac{d}{dx}\delta y\right)$$

Now, wouldnt it be nice if we could factor out that δy? but the 2^{nd} term has a $\frac{d}{dx}$ in front of it. What do do?

You've seen things like this before in freshman calculus. Recall integration by parts

$$\int_a^b udv = uv|_a^b - \int vdu$$

This method lets us take the "d" off of v and stick it onto u instead at the expense of the boundary term uv. Applying this idea to the problem at hand,

$$\int \frac{\partial F}{\partial y'}\frac{d}{dx}\delta y$$

$$\int udv \to u = \frac{\delta F}{\partial y'} \quad dv = \frac{d}{dx}\delta y$$

$$\therefore \int \frac{\partial F}{\partial y'}\frac{d}{dx}\delta y = \frac{\partial F}{\partial y'}\delta y\Big|_a^b - \int \frac{d}{dx}\frac{\partial F}{\partial y'}\delta y$$

But look! We said there is no variation on the end points so that

$$\frac{\partial F}{\partial y'}\delta y\Big|_a^b = 0 \quad \text{since} \quad \delta y|_a^b = 0$$

$$\therefore \quad \delta L = \int dx \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}\right)\delta y$$

Now, we want this to be extremized just like we do in freshman calculus. We recall that the extrema is where a derivative is zero or where the function doesn't change (linearly) as the variable changes, i.e. $df = 0$.

Here we want the "Functional Derivative" to be zero. Since δy is arbitrary the only way for δL to be 0 is if the integrand is 0.

$$\frac{\delta L}{\delta y} = 0 \Rightarrow \frac{\partial F}{\partial y} = \frac{d}{dx}\frac{\partial F}{\partial y'}$$

This is one of the most important equations in all life: the celebrated Euler-Lagrange equations!

$$\frac{d}{dx}\frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y} = 0$$

Solutions of this equation are precisely those which extremize our functional L. We have also inadvertently defined a new type of derivative, the functional derivative which become really useful in quantum field theory and stochastic calculus.

Now, in our example we have,

$$F = \sqrt{1 + y'^2}$$

Pro-tip: We could equally well extremize L^2!

$$L^2 = \int dx \left(1 + y'^2\right)$$

$\therefore$ we find derivatives of $F = \left(1 + y^2\right)$

$$\frac{\partial}{\partial y}\left(1+y'^2\right)=0$$
$$\frac{\partial}{\partial y'}\left(1+y'^2\right)=2y'$$
$$\frac{d}{dx}\frac{\partial}{\partial y'}\left(1+y'^2\right)=2y''$$

$\therefore$ we find the equation $2y''=0$

But, dear reader, you know that this is solved by equations of the form

$$y=mx+b$$

Duh: straight lines minimize distance in flat Euclidean space!
SUCH A BEAUTIFUL METHOD!

Can we do the same thing in physics with our action functional?
i.e. if $\delta S=0 \overset{?}{\Rightarrow} q(t)$ is the real life trajectory of our object? That is, the ones we find from Newton and Hamilton.

Let's try with $\mathscr{L}=KE-PE$

For an object moving under no forces we have

$$\mathscr{L}=\frac{1}{2}m\dot{q}^2$$

and

$$S=\int dt\frac{1}{2}m\dot{q}^2$$

$$\delta S=0\Rightarrow\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{q}}-\frac{\partial\mathscr{L}}{\partial q}=0$$
$$\frac{\partial\mathscr{L}}{\partial\dot{q}}=m\dot{q}$$
$$\frac{\partial\mathscr{L}}{\partial q}=0$$

so that

$$m\ddot{q}=0$$

but we've seen this equation before!

$$\sum F=ma$$

no forces $\Rightarrow ma = 0$

$$\Rightarrow x(t) = x_i + v_i t$$

straight line motion! Just what Newton and Galileo told us so long ago!

What about motion near the earth?

then $PE = mgx$

$$\therefore \quad \mathscr{L} = \frac{1}{2}m\dot{x}^2 - mgx$$

$$\begin{cases} \frac{d}{dt}\frac{\partial y}{\partial \dot{x}} = \frac{d}{dt}m\dot{x} = m\ddot{x} \\ \frac{\partial y}{\partial x} = -mg \\ y \Rightarrow m\ddot{x} - my = 0 \end{cases}$$

or $m\ddot{x} = mg$ or $\ddot{x} = g$

$$\Rightarrow a = g \ !$$

It works! Object fall with acceleration $g = -10\frac{m}{s^2}$ and follow Galileo's trajectories!

What about a spring?

$$\mathscr{L} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2$$

$$\frac{d}{dt}\frac{\partial y}{\partial \dot{x}} = m\dot{x} \quad \frac{\partial y}{\partial x} = -m\omega^2 x$$

$$m\ddot{x} = m\omega^2 x = m\left(\sqrt{\frac{k}{m}}\right)^2 x = kx$$

Dear reader, look at how much physics we have solved in just a few algorithmic lines of effort. No thinking involved. Just write things down. Energies are a LOT easier to define than forces! Now, these are all easy enough to solve using Newton's laws. What do we gain here?

What we gain is generality!

Suppose we want to find the equation of motion for a large rubber sheet.

We could model this as a bunch of springs connecting masses in a lattice and then take the continuum limit.

$\phi(x, y)$ represents the height of the sheet from equilibrium at point x, y We can show that

$$\mathscr{L} = \frac{1}{2}m\dot{\phi}^2(x,y) - \frac{1}{2}m\omega^2\phi^2(x,y)$$
$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{\phi}} = m\dot{\phi}(x,y) \quad \frac{\partial\mathscr{L}}{\partial\phi} = -m\omega^2\phi(x,y)$$
$$\therefore \quad m\ddot{\phi} + m\omega^2\phi = 0$$
$$\text{or} \quad \ddot{\phi} + \omega^2\phi = 0$$

This is called a "scalar field equation and it's solutions describe a vibrating membrane in space and time! This is starting to look like Quantum Field Theory so let's step back from a gorgeous subject.

All of this is called "Hamilton's Principle" or "the action principle" and still other names.

It states that "Nature determines the trajectory of a system by "extremizing" its action.

Subtly we say "extremize" rather than "minimize" - oddly we don't need to know whether action is maximum or minimum, not whether $\delta^2 S > 0$ or $\delta^2 S < 0$, we just need to know

$$\delta S = 0$$

Nobody knows "Why" this is true. Such a simple expression determines the evolution of almost everything in physics from gravity to Quantum mechanics

An interesting interlude

Minimum distances in arbitrary geometries.

We previously found the shortest path between two points in flat space is given by a straight line. What about on a sphere or some arbitrarily curved space?
On flat space we had $ds^2 = dx^2 + dy^2 + dz^2$ via Pythagoras.
If we generalize to curved spaces we will have some complicated expressions like

$$\begin{aligned} ds^2 =& xdx^2 + 2y^2dxdy + ydzdx \\ &+ \sin(xy)dz \end{aligned}$$

or something.
We can write the most general case as follows

$$ds^2 = g_{ij}dx^i dx^j$$

where we sum over i and j in $\in \{x, y, z\}$ e.g. $ds^2 = g_{xx}dx^2 + g_{xy}dxdy + g_{yx}dydx + g_{yy}dy^2$ for two dimensional curved space.
here g_{ij} are functions in general
$g_{ij} = g_{ji}$ is called "The Metric"
and completely encodes the geometry of a general space.
We can write g_{ij} is a matrix

$$G = \begin{bmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{bmatrix}$$

For Euclidean space $g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \delta_{ij}$
Viz,

$$g_{ij}dx^i dx^j = dx^2 + dy^2$$

For an arbitrary space the total length of a curve is given by

$$L = \int ds = \int_A^B \sqrt{g_{ij}dx^i dx^j}$$

we can parameterize the curve by t, λ, l or whatever variable is convenient such that

$$L = \int \frac{d\lambda}{d\lambda} ds \int \frac{ds}{d\lambda} d\lambda = \int_{\lambda_A}^{\lambda_B} \sqrt{g_{ij}\frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}} d\lambda$$

The process is the same only m,uch more algebraically complicated.
Consider

$$L^2 = \int g_{ij}\dot{x}^i \dot{x}^j d\lambda$$

Where

$$\dot{x}^i = \frac{dx^i}{d\lambda}$$

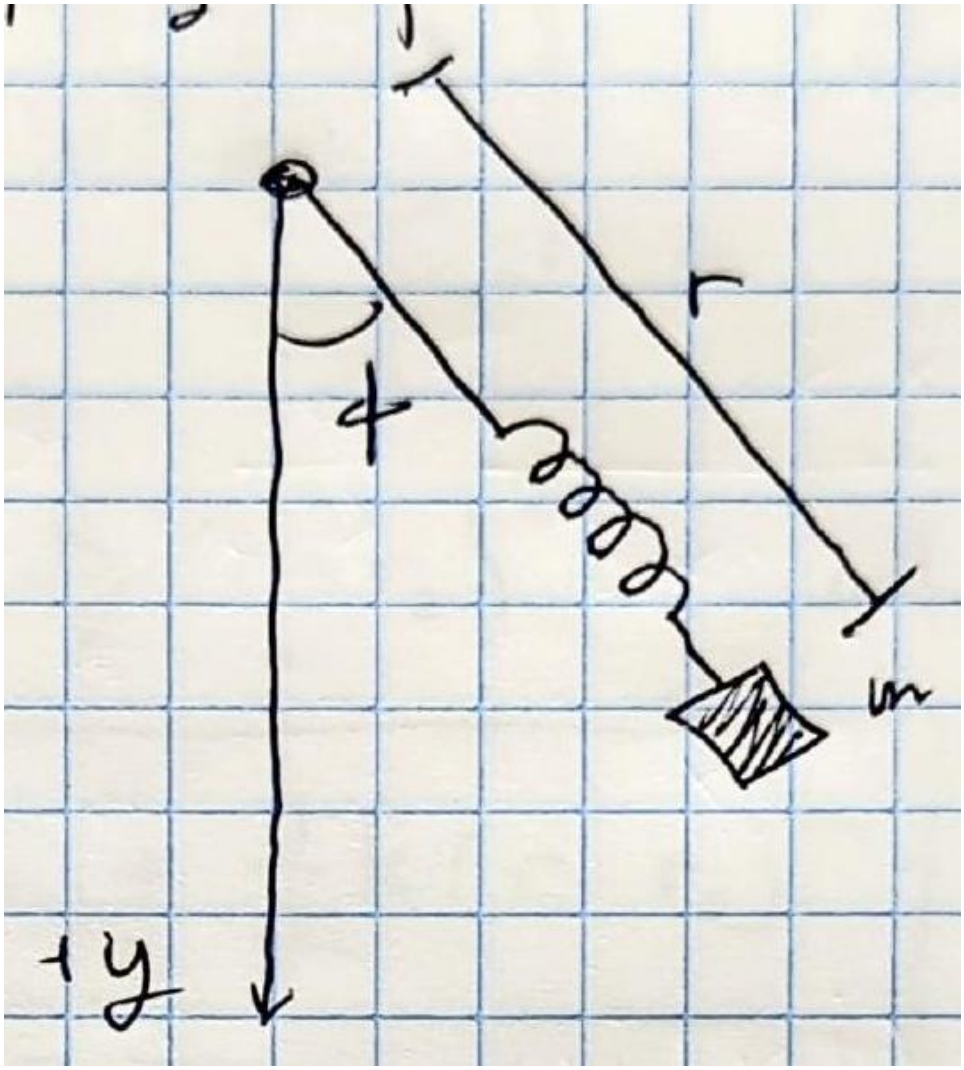
then $\delta L^2 = 0 \Rightarrow \frac{d}{d\lambda}\frac{\partial g_{ij}\dot{x}^i\dot{x}^j}{\partial \dot{x}^i} - \frac{\partial g_{ij}\dot{x}^i\dot{x}^j}{\partial x^i} = 0$ now $g_{ij} = g_{ij}(x)$ but not $\dot{x}$ so we will need to take care with the derivatives:

$$\frac{d}{d\lambda}2g_{ij}\dot{x}^i - \frac{\partial}{\partial x^k}\left(g_{ij}(x)\dot{x}^i\dot{x}^j\right)$$
$$= 2g_{ij}\ddot{x}^i - \frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j - g_{ij}\frac{\partial \dot{x}^i}{\partial x^k}x^j - g_{ij}\dot{x}^i\frac{\partial \dot{x}_j}{\partial x^k}$$
$$\frac{\partial \frac{d}{d\lambda}x^i}{\partial x^k} = \frac{d}{d\lambda}\frac{\partial x^i}{\partial x^k} = \frac{d}{d\lambda}\delta^i_k = 0$$
$$\ddot{x}^k - \frac{1}{2}\Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$$

is called the geodesic equation and it's solutions produce the shortest distance curves between two points in an arbitrarily curved space.
$g_{ij} = \delta_{ij}$ for flat space. Therefore, $\ddot{x}^i = 0 \Rightarrow$ straight lines as we previously found.

Physics Example:

Find the equation of motion EOM of a mass/spring pendulum near the Earth. see below



(1) Since we have a dynamic angle ϕ and dynamic length r polar coordinates are appropriate.
take

$$y = r\cos\phi$$
$$x = r\sin\phi$$

then

$$dy = dr\cos\phi + rd\phi\sin\phi$$
$$dx = dr\sin\phi - rd\phi\cos\phi$$

$$\dot{y} = \dot{r}\cos\phi + r\dot{\phi}\sin\phi$$

$$\dot{x} = \dot{r}\sin\phi - r\dot{\phi}\cos\phi$$

$$

$$\mathscr{L} = KE - PE$$

$$KE = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2$$

$$= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right)$$

$$PE = -mgy + \frac{k}{2}\left(r - r_0\right)^2$$

where r_0 is the natural length of the spring.
Now, we have two E-L equations; one for each coordinate.
(1)

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{r}} - \frac{\partial\mathscr{L}}{\partial r} = 0$$

(2)

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{\phi}} - \frac{\partial\mathscr{L}}{\partial\phi} = 0$$

$$\therefore \quad \frac{d}{dt}\left(mr^2\dot{\phi}\right) = -mgr\sin\phi$$

from (1) (2) we have

$$mr\ddot{\phi} = -mg\sin\phi - 2m\dot{r}\dot{\phi}$$

$$m\ddot{r} = mr^2\dot{\phi}^2 + mg\cos\phi - k\left(r - r_0\right)$$

these two differential equations yield the solutions for $r(t)$ and $\phi(t)$. Try to solve these using a good guess or some elaborate method from your studies of differential equations. For now, id just like you to know about this particular method in mechanics!

Now for the quantum harmonic oscillator: one of the most important moments in your life!

THE QUANTUM HARMONIC OSCILLATOR

Quantum Approach

In the Quantum world we usually follow Hamilton's approach but there are many other methods. Our quantum Hamiltonian approach is motivated by our realization that the "Hamiltonian Operator Generates time evolution of quantum states".

e.g. we previously found that

$$|\psi(t)\rangle = e^{\frac{i}{\hbar}\hat{H}t}|\psi(0)\rangle$$

and we had decided that it's usually a great idea to use "energy eigenstates" with which to expand our general $|\psi\rangle$ in terms of. So we seek the eigenvalues and eigenvectors of

$$\hat{H}|E_n\rangle = E_n|E_n\rangle$$

and then a general $|\psi\rangle$ will be a linear combination of these $|E_n\rangle$

$$|\psi\rangle = \sum c_n|E_n\rangle$$

where $c_n \in \mathbb{C}$ are the probability amplitudes for state $|E_n\rangle$

Now, at this point we need to write $\hat{H}$ explicitly if we hope to find a solution to any given system.

It can be "shown"(im unsure if it's been proven? physically nature seems to like it at least), but for now trust, that $\hat{H}$ "has the same form" as our classical $H(p,x)$ except everywhere we have x, p we turn them into operators $\hat{x}, \hat{p}$. Mysterious! We shall take a different approach later where everything involves classical variables and no operators! However, this is the typical, or "canonical" route towards quantum physics.

For the SHO

$$H(p,x) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \to \hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$$

We simply placed little hats on top of our dynamical variables!

Here we have two routes to finding $|E_n\rangle$

1. Differential equation. In the position basis we have

$$\hat{p} \to -i\hbar \frac{d}{dx} \quad \hat{x} \to x$$

recall that we previously found $\hat{p}$ in terms of position as the generator of spatial translations. This yields a differential equation that can, in principle, be solved for $\psi(x) = \langle x|\psi\rangle$

A better, and MUCH more clever, route is

2. Algebraically - this leads to absolutely PROFOUND insights due to Dirac and forms the foundation of all modern physics. It is not an obvious thing to do. Dirac was an incredible genius full of insight and unafraid to try even the most ridiculous procedures! It boggles my mind that this was found! We shall dutifully follow Dirac and we will learn a LOT.

Algebraic Solution

The clever thing Dirac tried was to notice the high school fact that (Dirac used this idea in other profound areas as well!)

$$x^2 + y^2 = (x + iy)(x - iy)$$

In our SHO we have

$$\tilde{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}^2$$

So perhaps we can "factor" the Hamiltonian operator. Maybe this will lead somewhere; maybe it's a dead end. With new discoveries nobody knows where a path will lead. It's up to the bold (and often young!) to adventure down curious paths!

Here we can algebraically move terms around in our $\hat{H}$ to make everything dimensionless and then factor the operator by defining operators

$$a = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\sqrt{\frac{1}{2m\omega\hbar}}\hat{p}$$
$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - i\sqrt{\frac{1}{2m\omega\hbar}}\hat{p}$$

This is of the same form as $x^2 + y^2 = (x + iy)(x - iy)$ but with a bunch of constants. Factor H yourself, youll get a lot of experience by doing so.

NOTE This works only because $\hat{x}, \hat{p}$ are Hermitian operators!
viz, $\hat{x} = \hat{x}^\dagger \quad \tilde{p} = \hat{p}^\dagger$

Now we need another ingredient. The commutator. This follows directly from our previous knowledge that momentum "generates" spatial transitions

You can show that (DO IT)

$$[\hat{x}, \hat{p}] = i\hbar\mathbb{1}$$

recall $[A, B] = AB - BA$ is called the "commutator"

Often you'll see $[\hat{x}, \hat{p}] = i\hbar$ with the understanding that there's really an $\mathbb{1}$: the identity operator.

In the position basis $\hat{p} \to -i\hbar\frac{d}{dx}$ and $\hat{x} \to x$

$$[\hat{x}, \hat{p}] = \left[x, -i\hbar\frac{d}{dx}\right] = \left(-i\hbar x\frac{d}{dx} + i\hbar\frac{d}{dx}(x)\right)$$

When analyzing commutators the operators need vectors/functions to act upon, so pick a general $f(x)$

$$[\hat{x}, \hat{p}]f(x) = (-i\hbar x\frac{d}{dx}f(x) + i\hbar\frac{d}{dx}(xf(x))) = i\hbar f(x)$$

It works!!

Back to our $\hat{a} : \hat{a}^\dagger$
if $[\hat{x}, \hat{p}] = i\hbar\mathbb{1}$ then you can show (DO IT!)

$$\left[a, a^\dagger\right] = \mathbb{1}$$

by plugging our definitions of $\hat{a}, \hat{a}^\dagger$

Now, since

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\left(a + a^+\right)$$
$$\hat{p} = -i\sqrt{\frac{m\omega\hbar}{2}}\left(a - a^+\right)$$

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \quad \text{becomes}$$
$$\hat{H} = \frac{1}{4}\hbar\omega\left[-\left(a - a^\dagger\right)^2 + \left(a + a^\dagger\right)^2\right]$$
$$= \frac{1}{2}\hbar\omega\left(aa^\dagger + a^\dagger a\right)$$

but $\left[a, a^\dagger\right] = 1$ means $aa^\dagger - a^\dagger a = 1$

$$\therefore \quad \Rightarrow aa^\dagger = \mathbb{1} + a^\dagger a$$

and

$$\hat{H} = \hbar\omega(a^\dagger a + \frac{1}{2})$$

This is one of the most important equations in all Humanity! We have reached a profound summit of human knowledge!!!

But what the heck do we do with this?

Recall our primary goal: find eigenvalues and eigenvectors of $\hat{H}$. We shall do so by using this form.

We know that these should be energy states!
viz

$$\hat{H}\,|E_n\rangle = E_n\,|E_n\rangle$$

Let's call $a^\dagger a = \hat{N}$
then we'd love to know the eigenvectors and eigenvalues of $\hat{N}$: everything else is a number. If we knew these then we'd have our energy states.

So we wish to solve:

$$\hat{N}|n\rangle = n|n\rangle$$

But we said $|E_n\rangle$ are the eigenstates of $\hat{H}$ and everything else is a number, so we find that $|E_n\rangle = |n\rangle$!! So let's relabel what we used to call $|E_n\rangle$ and call it $|n\rangle$

We have

$$\hbar\omega(\hat{N} + \frac{1}{2})|n\rangle = E_n|n\rangle$$

Then since $\hat{N}|n\rangle = n|n\rangle$ we find

$$E_n = \hbar\omega(n + \frac{1}{2}) \quad n \in \mathbb{N}$$

BOOM! Those are the energy states of the quantum harmonic oscillator!! We find each energy "level" is equally spaced and n ranges from 0 to ∞. The $n = 0$ state is called the "ground state" or the "vacuum state". Notice that the energy of this ground state is non-zero. Evidently this system always has some energy. But this is okay because usually we care about "changes" in energy rather than energy. The only place where this becomes problematic is in Einstein's General Relativity where space-time responds to Energy rather

than energy changes. We won't concern ourselves with this important feature of physics but do keep it in the back of your mind.

Now we wish to study how these a and $a^\dagger$ operators act on $|n\rangle$.

The following is NOT an obvious thing to do so bear with it and follow along for the important punchline.

We need the following result to proceed

$$[AB, C] = ABC - CAB = A[B, C] + [A, C]B$$

Try to prove this (remember, order matters! These are operators/matrices)

Now,

$$[N, a] = [a^\dagger a, a] = -a$$

$$[N, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger$$

Let's consider

$$\hat{N}a|n\rangle$$

and attempt to massage this into something like a $a\hat{N}|n\rangle$ because we know $N|n\rangle = n|n\rangle$.

We use the commutators above

$$\begin{aligned} \hat{N}a|n\rangle = (a\hat{N} - a)|n\rangle &= a(\hat{N} - 1)|n\rangle \\ &= a(n-1)|n\rangle \\ &= (n-1)a|n\rangle \end{aligned}$$

But $(n-1)$ is just a number

$$\hat{N}a|n\rangle = (n-1)a|n\rangle$$

We stare at this and after awhile a thought occurs to us. This says that the vector $a|n\rangle =$, say $|\alpha\rangle$ is ALSO an eigenvector of $\hat{N}$! Its eigenvalue is $(n-1)$!

$$\hat{N}|\alpha\rangle = (n-1)|\alpha\rangle$$

The eigenvector that has eigenvalue $(n-1)$ is simply $|n-1\rangle$
evidently $\hat{a}|n\rangle \to |n-1\rangle$

a changes the state $|n\rangle$ so as to "reduce" n by 1 . We have no idea what $|n\rangle$ is or what it means or even how to express it. Yet we KNOW a HAS to

act this way!
Similarly, you can now show that

$$\hat{N}a^\dagger|n\rangle = (n+1)a^\dagger|n\rangle$$

We have found that:

$$a|n\rangle \to |n-1\rangle$$

and

$$a^\dagger|n\rangle \to |n+1\rangle$$

Now, we dont yet know what the full affect of these operators are on $|n\rangle$ yet.
we have,

$$a^\dagger|n\rangle = c_n|n+1\rangle$$

and

$$a|n\rangle = b_n|n-1\rangle$$

We wish to find c_n and b_n, then we would be done.
The last piece comes from

$$\langle n|\hat{N}|u\rangle = \langle n|n|n\rangle = n\langle n|n\rangle = n$$

However we want $\langle n|n\rangle = 1$. That is to say, these states should be normalized. Furthermore we know $\hat{N} = a^\dagger a$

$$\begin{aligned} \therefore \quad & \langle n|a = \left[a^\dagger|n\rangle\right]^\dagger \\ & \langle n|aa^\dagger|n\rangle = c_n^2 = n+1 \\ \therefore \quad & a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \qquad \text{(iii)} \\ \text{and} \quad & a|n\rangle = \sqrt{n}|n-1\rangle \end{aligned}$$

Now we have everything!

$$\begin{aligned} & \hat{H}|n\rangle = E_n|n\rangle \\ & \hbar\omega(\hat{N}+1/2)|n\rangle = \hbar\omega(n+1/2)|n\rangle \\ & E_n = \hbar\omega(n+1/2) \end{aligned}$$

The ground state is $|0\rangle$
From this state we can build all other by acting with $a^\dagger$ repeatedly
e.g.

$$a^\dagger|0\rangle = \sqrt{1}|1\rangle$$

$$a^\dagger a^\dagger|0\rangle = \sqrt{1}a^\dagger|1\rangle = \sqrt{2}|2\rangle$$

and so on. We simply climb the ladder to which ever state we want to visit. In general,

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle$$

Similarly a let's us climb down the ladder

$$a \quad \text{is called the annihilation operator}$$

$$a^\dagger \quad \text{is called the creation operator}$$

That's "all" there is to it!. We have officially solved the quantum harmonic oscillator and there wasn't ANY calculus or differential equations involved! Now, if, for some reason, we wish to know what $\psi_n(x)$ is in terms of position then we'll generally have to solve differential equations. In particular for $|\psi\rangle \in \mathcal{H}$

$$|\psi\rangle = \sum \alpha_n|n\rangle$$

and

$$\psi(x) = \langle x|\psi\rangle = \langle x|(\sum \alpha_n|n\rangle)$$

$$= \sum \alpha_n\langle x|n\rangle$$

Therefore, in order to find a general $\psi(x)$ we must figure out how to compute $\langle x|n\rangle$

For the ground state in terms of $\psi(x)$ we have

$$\psi(x) = \langle x|0\rangle$$

Whichever way we may wish to proceed towards $\psi(x)$ we still have to solve differential equations. I merely quote this difficult solution as

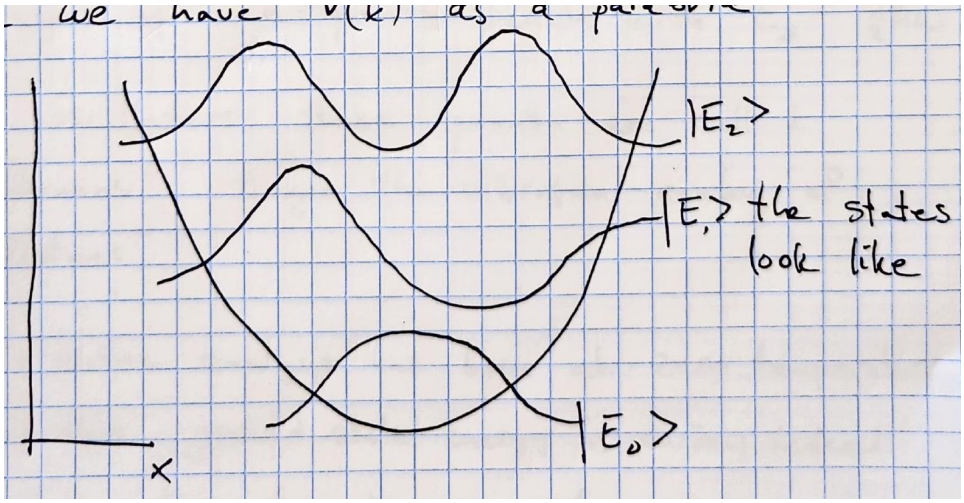
$$\psi_n(x) = \langle x|n\rangle = \sqrt{\frac{m\omega}{\hbar\pi^{1/2}2^n n!}}H_n(\frac{xm\omega}{\hbar})e^{\frac{1}{2\hbar^2}m^2x^2}$$

where $H_n(z)$ is a Hermite polynomial of degree n. Theyre special functions that often must be looked up. Let's check $\psi_0(x)$, by looking up $H_0(z) = 1$

$$\psi_0(x) = \sqrt{\frac{m\omega}{\hbar\pi^{1/2}}} e^{\frac{1}{2\hbar^2} m^2 x^2}$$

This is a standard Gaussian. The ground state looks like a bell curve as a function of space.

The first few states look like:



The energy levels are all evenly spaced. The parabola is a sketch of the spring potential.

The obvious interpretation is that n labels the state of a single object and how "excited" it is. This is the first interpretation we are taught but there is a DEEPER and more clever interpretation,

The deeper interpretation is that n labels "the number of quanta". A single quantum has energy $E_0 = \frac{1}{2}\hbar\omega$. So, for example, the state $n = 27$ we say "there are 27 particles in our system" .

E_n looks like the total energy of n-quanta each with $E_0 = \frac{1}{2}\hbar\omega$ energy!

For masses connected by spring-like potentials these quanta are called "phonons". They are the vibrational-analog of quanta of light called photons.

Now, this entire analysis has been at zero temperature. This ground state energy is then curious indeed! $T = 0$ does not imply zero motion or energy! We can measure the position of our zero T phonon and we WILL find it moving about! Weird!

This example exemplifies how we typically proceed in solving a quantum system. In general we do the following:

1. We write down the Classical Hamiltonian $H(p, x)$ where

$$H \equiv \text{ Kinetic } + \text{ Potential}$$

2. We promote all the x's and p's into operators. They get little hats:

$$p \to \hat{p}$$
$$q \to \hat{q}$$

3. We demand our operators to satisfy

$$[\hat{x}, \hat{p}] = i\hbar\mathbb{1}$$

CAUTION if your classical $H(x, p)$ has a term like $2xp$ then you DO NOT KNOW the "correct" order quantum mechanically. Remember order matters in our quantum theory but classically we couldve also written $2px$. We can get around this by

$$2xp \to xp + px$$

That is to say, we force it "by hand". Practically this doesn't seem to be an issue in everyday computations but philosophically it is very ad hoc and doesnt have a firm physical justification. It often doesnt come up in typical physical system but if it does, this sneaky trick, usually works out okay.

4. Find a symmetry of your system. If $\hat{H}$ does not explicitly depend an time, then you almost always want to choose to use energy states as your basis. i.e.

$$\hat{H} \left|E_n\right\rangle = E_n \left|E_n\right\rangle$$

If there is no $\hat{x}$ operator in your Hamiltonian then spatial translation is a symmetry and you may want to solve it using momentum states.
e.g.

$$\hat{H} = \frac{\hat{p}^2}{2m}$$

for a free particle

$$\hat{H}|p\rangle = \hat{p}|p\rangle \Rightarrow E = \frac{p^2}{2m}$$

The system under study usually hints at what basis states to choose but experiment is king. If the only thing you can detect is position in your lab then you'll want to use $|x\rangle$ position states which may not be a symmetry of your Hamiltonian.

5. Clever Diagonalization - Solving

This is an art. It will not be obvious but if you can do it for some complex system then you usually become famous like Dirac, Bogoliubov, is Heisenberg! In the SHO we did this with
$\hat{N} = a^\dagger a$ to find states $|n\rangle = \left(a^\dagger\right)^n |0\rangle$ and $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$

That's it! This process is called "Canonical Quantization" and, as the name implies, was the original method used to solve quantum systems. The reason why taking $x \to \hat{x}$ and $p \to \hat{p}$ works is a complex mathematical and physical journey and at this level we will simply trust it works because experiments seem to agree.

Down a fruitful path?
You might naturally ask "What are the eigenvectors and eigenvalues of a and $a^\dagger$?"

They aren't $|n\rangle$ because $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$. We don't get $|n\rangle$ back after $a^\dagger$

The eigenvalues and eigenvectors of a, $a^\dagger$ are called "coherent states" and are extremely important in the study of Quantum Optics.
viz

$$a|z\rangle = z|z\rangle$$

$|z\rangle$ are called coherent states and are given by

$$|z\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}|n\rangle$$

where z is the complex eigenvalue of a.

This coherent state says that $|z\rangle$ is a linear combination of ALL values of $|n\rangle$ i.e. of ALL energy levels or phonons.
i.e.

$$|z\rangle \sim \text{ (no photons)} + \text{(1 phonon)} + \text{(2 phonons)} + \ldots$$

Weird! If you know $|z\rangle$ then you don't have any knowledge of n or the actual energy of your system. In optics these coherent states behave almost classically which leads them to be quite useful but I shall leave this mysterious statement unaddressed for now.

Now let's turn our attention towards another quantum problem which we shall use a differential equation to solve. This is the problem of a quantum object trapped inside a box bounded by an infinite potential.

Example Computation

Infinite Potential Well or "a quantum particle in a box". We consider the following potential energy of our quantum system. Our object of interest is $|\psi\rangle$ which is confined to a box with zero possibility of escape. We will ask the question: what is $\psi(x)$? In particular we want to find the projection of $|\psi\rangle$ in the position basis. i.e. $\langle x|\psi\rangle = \psi(x)$.

We know that,

$$V(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq a \\ \infty & \text{elsewhere} \end{cases}$$

Outside of the box we assume $\psi(x) = 0$. For a finite potential well this may not be true. Later we will see that for finite potentials the quantum object can "leak out" of the box.

Following our guidelines we write down the Hamiltonian and promote dynamical variable to operators

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$$

We wish to solve

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

Using the position basis we recall that $\hat{p} = -i\hbar\frac{d}{dx}$.

Inside the box $V = 0$ so we find the differential equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi \quad 0 \leq x \leq a$$

re-arranging

$$\frac{d^2\psi}{dx^2} = -k^2\psi$$

$$k = \frac{\sqrt{2mE}}{\hbar}$$

or

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0$$

Dear reader, does this equation look familiar?!

It looks like a spring system from classical physics only instead of time our spring oscillates in space! We know how to solve this. We simply try oscillatory solutions in space. We guess,

$$\psi(x) = Ae^{\pm ikx}$$

or

$$\psi(x) = A\cos(kx) + B\sin(kx)$$

Now, we expect $\psi(0) = \psi(a) = 0$ because the walls are infinite.

$$\therefore \quad \psi(0) = A\sin 0 + B\cos 0 = B \Rightarrow B = 0$$

We find that $\psi(x)$ must be a pure sine wave!
Now,

$$\psi(a) = 0 \Rightarrow A\sin ka = 0$$

This says that either $A = 0$ or $\sin(ka) = 0$

If $A = 0$ then we simply have nothing in the box. Boring! So we take $\sin(ka) = 0$.

$$\sin(ka) = 0 \Rightarrow ka = 0, \pm\pi, \pm 2\pi, \cdots$$

or $k = \frac{n\pi}{a}$ for $n \in \mathbb{N}$

We get MANY discrete solutions. Let's relabel k as k_n

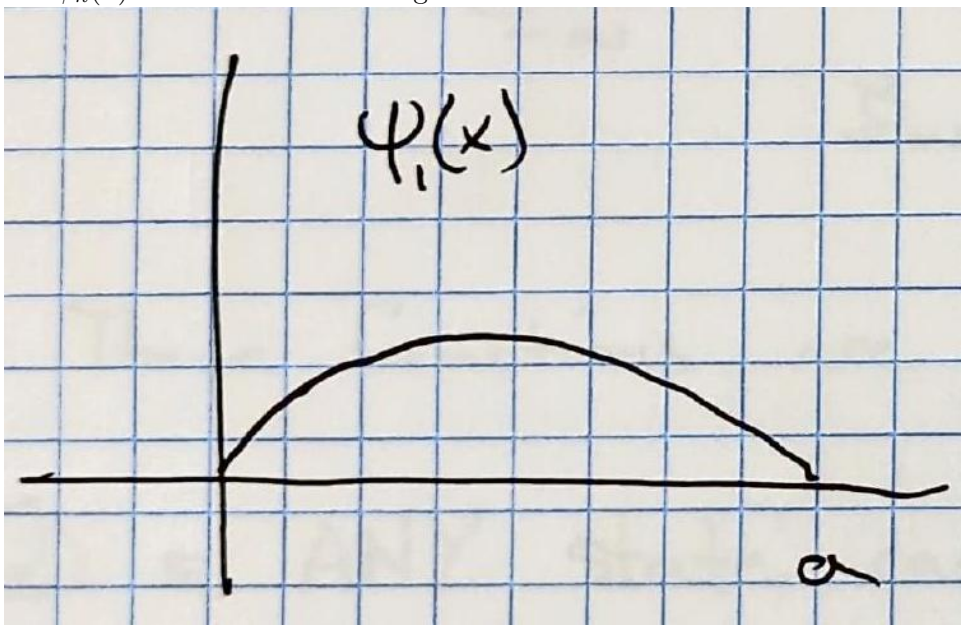
$$\therefore \quad \psi(x) = A\sin\left(\frac{n\pi}{a}x\right)$$

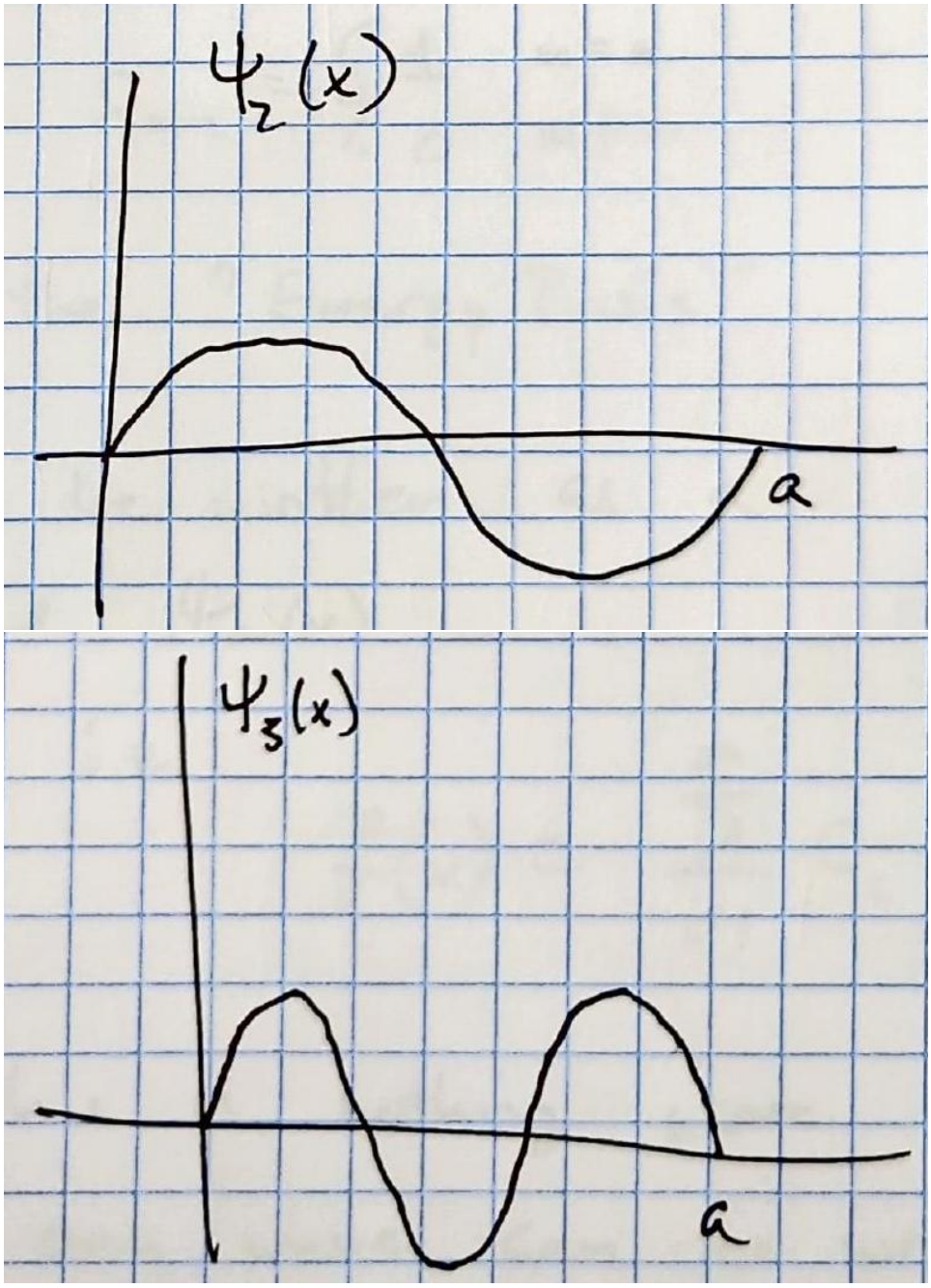
But how do we find the amplitude A? THINK! The particle is in the box somewhere. Therefore we have a 100% chance of finding it inside if we were to look.

$$\therefore \int_0^a p(x)dx = 1 = \int_0^a \psi^*(x)\psi(x)dx = 1$$
$$= \int_0^a A^2 \sin^2(k_n x)\, dx = A^2 \frac{a}{2} = 1$$
$$\Rightarrow A = \sqrt{\frac{2}{a}}$$

How nice! We used the boundaries and the fact that probabilities sum to 1 to completely determine our solution. But we also know that $k_n = \frac{\sqrt{2mE_n}}{\hbar}$. That's what we called k to begin with. Therefore we can find all the different possible energies of this object in a box!

We find there are MANY different $\psi(x)$ each with a different value of k or n and E_n. We shall label them $\psi_n(x)$
our $\psi_n(x)$ looks like the following





and so on!

As E_n gets larger the $\psi_n(x)$ gets more wiggly. This is a general feature of quantum objects.

We say that these $\psi_n(x)$ are the "energy basis" states of this "infinite well potential" system.

Notice the following curious facts:

1.

$$\int_0^a \psi_m^*(x)\psi_n(x)dx = 0$$

$$m \neq n$$

We say that $\psi_m(x)$ and $\psi_n(x)$ are orthogonal just like if $\vec{a} \cdot \vec{b} = 0$

In general "basis functions" satisfy

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

$$\delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

2. ANY state can be written as a linear combination of $\psi_n(x)$

$$f(x) = \sum_{i=1}^{\infty} c_i\psi_i(x) = \sqrt{\frac{2}{a}} \sum_i c_i \sin\left(\frac{n\pi x}{a}\right)$$

This is called a Fourier Series. Any arbitrary wave can be written as a sum of sines/cosines.

The lesson here is often we get discrete Energy spectra from solving Schrödinger's equation - a continuous differential equation.

Recall that we argued that the full quantum state as a function of space AND time is $\Psi(x,t) = \psi(x)\phi(t)$ where $\phi(t) = e^{-\frac{iE_nt}{\hbar}}$

So our FULL solution, which describes ANY possible state of our object in this potential well is

$$\Psi(x,t) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi x}{a}\right) e^{-\frac{iE_nt}{\hbar}}$$

where

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$$

or more abstractly we can write

$$|\Psi\rangle = [c_1 |\psi_1\rangle + c_2 |\psi_2\rangle + \dots c_n |\psi_n\rangle] \, e^{-\frac{iE_n t}{\hbar}}$$

Generally the time dependence of a wavefunction describing a quantum system is almost always $e^{-\frac{iE_n t}{\hbar}}$. We say the complex phase of the quantum system oscillates in time.

Our full and most general solution says "before we measure the energy of the object it is in the arbitrary state $\Psi(x,t)$"

If we measure the energy using the energy operator it will FORCE Ψ to take on one and only one of the specific energy states $|\psi_n\rangle$ with probability given by c_n^2. We can only ever find our object in the box with one definite energy.

Now, we can ask fun questions such as, "what is the average position of my object if it's in some state?" Or, "if I look where am I likely to find it?"

The proper way to proceed is to find that average value of our $\hat{x}$ operator. Way way back we defined the average of an operator as

$$\langle\psi|\hat{T}|\psi\rangle = \int \langle\psi|x\rangle\langle x|\hat{T}|x\rangle\langle x|\psi\rangle = \int \psi^*(x) T \psi(x) dx$$

Here we know $\psi(x)$ and the form of our x operator in the position basis, therefore

$$\langle\hat{x}\rangle = \int \psi^*(x) x \psi(x) dx$$

$$= \int x\psi^*(x)\psi(x)dx = \int xP(x)dx$$

where $P(x) = |\psi(x)|^2$ is the probability density of finding x.

This formula is straight out of something you'd see in a textbook on statistics where $\langle\hat{x}\rangle$ is the mean value of a probability distribution. By golly, that's what we've got here. So for any state $\psi_n(x)$ we can find the average position of our object. We immediately see that since our $\psi_n(x)$ are sine waves then at the nodes of our sine function there will be ZERO chance of finding our object.

This is baffling - the object is buzzing around in there but it'll never be found at certain places? Well, sure. But this quantum thing is like NOTHING you're familiar with in everyday life. Even though we like to imagine electrons as tiny particles what matters most is this $\psi_n(x)$ wave, and sure enough from some electron in a box there are places for it to never be found.

Your task: find the average position of the object in the energy state $\psi_4(x)$.

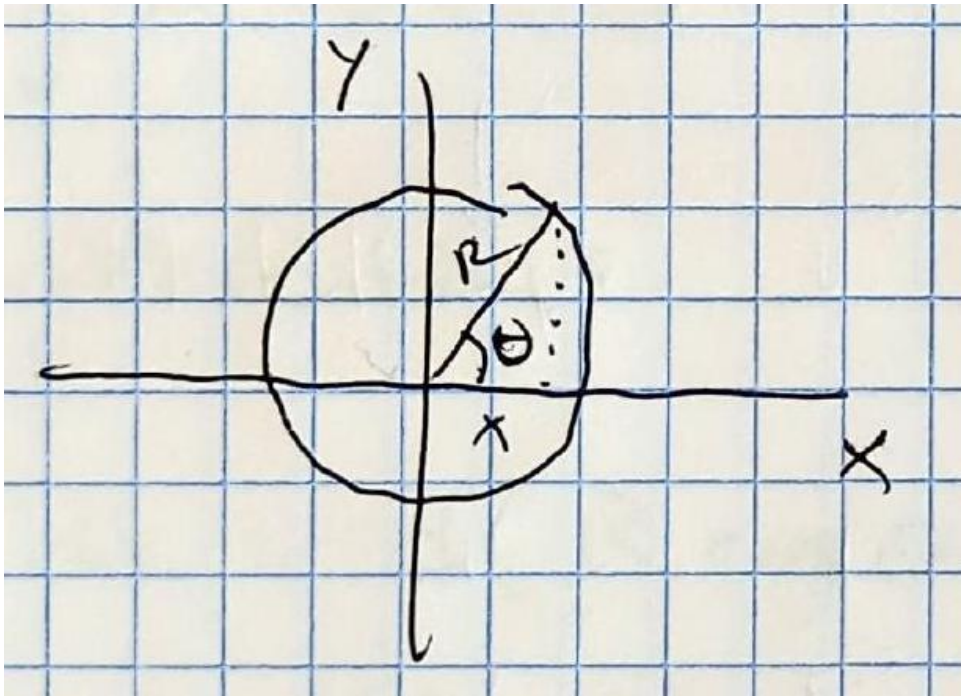
Now, think about what happens as the energy increases. Since E_n goes like n^2 then the larger E_n is the larger n is. But larger n means our $\psi_n(x)$ gets more "wiggly" as the energy and n increases. But, the number of nodes and peaks of our $\psi_n(x)$ increases as E_n increases. So perhaps as $n \to \infty$ the probability of finding a high energy object at position x becomes uniform and therefore we can expect to find the object anywhere in our box with equal probability. Or isit the other way around? As you increase energy the number of nodes goes to infinity and then you have zero chance of finding the object anywhere! Your task: Show which is correct by taking the limit of $|\psi_n(x)|^2$.

Fourier and a Circular Universe

Consider a free quantum object on a circular universe of radius R
We impose two requirements in this universe:

1. $\psi(\theta) = \psi(\theta + 2\pi n) \quad n \in \mathbb{N}$
2. $\int_0^{2\pi} |\psi(\theta)|^2 d\theta = 1$

There are no interactions at all so all the energy is all kinetic



In cartesian coordinates our Hamiltonian is simply the sum of kinetic energies in the x and y directions

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} = -\frac{i\hbar}{2m}\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2}\right)$$

where we recall the momentum operator as the generator of translations in the position basis.

$$\text{viz} \quad \hat{p}_x \to -i\hbar\frac{d}{dx}$$

We naturally want everything written in terms of an angular coordinate θ. The system is ultimately one dimensional.

We seek a way to write these cartesian derivative operators $\frac{d}{dx}; \frac{d}{dy}$ in terms of polar coordinates (r, θ). For now we keep r arbitrary and later we will set it equal to R

We have the coordinate maps

$$x = r\cos\theta$$

$$y = r\sin\theta$$

To facilitate our goal consider an arbitrary function $f(x,y)$
We can use the chain rule to find the polar derivatives in terms of Cartesian derivatives:

$$\frac{df(x,y)}{d\theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}$$

$$\frac{df(x,y)}{dr} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}$$

but we already have

$$\frac{\partial x}{\partial \theta} = \frac{\partial}{\partial \theta} r\cos\theta = -r\sin\theta$$

$$\frac{\partial y}{\partial \theta} = \frac{\partial}{\partial \theta} r\sin\theta = r\cos\theta$$

$$\therefore \frac{df}{d\theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}$$

$$\therefore \quad \frac{d}{d\theta} = -r\cos\frac{\partial}{\partial x} + r\sin\theta\frac{\partial}{\partial y}$$

and similarly for $\frac{df}{dr}$

So we have $\frac{d}{d\theta}$ and $\frac{d}{dr}$ in terms Cartesian derivatives. Therefore we may invert this pair of expressions to arrive at

$$\frac{d}{dx} = \cos\theta\frac{d}{dr} - \frac{1}{r}\sin\theta\frac{d}{d\theta}$$
$$\frac{d}{dy} = \sin\theta\frac{d}{dr} + \frac{1}{r}\cos\theta\frac{d}{d\theta}$$

This looks almost like our $2D$ rotation matrix applied to a "vector" of derivative operators modulo the $\frac{1}{r}$ terms. At any rate we can now replace our Cartesian derivatives with their polar counterparts.

Now, after these generalities we will impose $r = R$ and $dr = dR = 0$. So that,

$$\frac{d}{dx} = -\frac{1}{R}\sin\theta\frac{d}{d\theta}$$

$$\frac{d}{dy} = \frac{1}{R}\cos\theta\frac{d}{d\theta}$$

then

$$\frac{d^2}{dx^2} = \frac{1}{R^2}\sin^2\theta\frac{d^2}{d\theta^2}$$
$$\frac{d^2}{dy^2} = \frac{1}{R^2}\cos^2\theta\frac{d^2}{d\theta^2}$$

Therefore

$$\hat{H} = -\frac{i\hbar}{2m}\frac{1}{R^2}\frac{d^2}{d\theta^2}$$

this $2D$ problem has simplified to a VERY simple $1D$ problem merely by choosing proper coordinates. This is a common theme in physics: choose the coordinate system that matches with the symmetries of your system!

Perhaps we can solve this Schrödinger Equation (recall that we separate the time dependent part of $\Psi(x,t)$ from the spatial part $= \Psi(x,t) = \psi(x)\phi(t)$) and solve

$$\hat{H}\psi(\theta) = E\psi(\theta)$$

There may be many solutions indexed by n for now we will just write $\psi(\theta)$ without the index and then reintroduce it when and where we have to. Plugging in, we have

$$-\frac{i\hbar}{2mR^2}\frac{d^2\psi(\theta)}{d\theta^2} = E\psi(\theta)$$

brushing aside all the constants for a moment we seek a $f(\theta)$ such that

$$\frac{d^2f}{d\theta^2} = Ef(\theta)$$

This is a standard separation of variables. Take

$$\frac{d^2f}{f} = Ed\theta^2$$

integrate twice and find (modulo constants)

$$f(\theta) \sim e^{\pm\sqrt{E}\theta}$$

as the solutions. With all of our constants properly put back in we have,

$$\psi(\theta) = Ae^{\pm i\frac{R}{\hbar}\sqrt{2mE}\theta} = A\xi$$

where ξ is just a placeholder for the exponential part used to save space in the following computation. This looks like a standard wave-like solution (sines and cosines) a la Euler's identity.

Now, we don't know what our integration constant or E are but we have two major pieces of information we have yet to leverage:

1. $\int_0^{2\pi} |\psi(\theta)|^2 d\theta = 1$
2. $\psi(\theta) = \psi(\theta + 2\pi n)$

Using (1)

$$\int_0^{2\pi} A^2 \xi^*(\theta)\xi(\theta) d\theta$$

but notice

$$\xi^*(\theta) = \xi^{-1}(\theta) \Rightarrow \xi^*(\theta)\xi(\theta) = 1$$

the exponentials perfectly cancel! So

$$\int_0^{2\pi} A^2 d\theta = 2\pi A^2 \quad \forall\theta$$

but the sum of probabilities must $= 1$ from (1). Therefore,

$$\Rightarrow \quad 2\pi A^2 = 1 \Rightarrow A = \frac{1}{\sqrt{2\pi}}$$

Now to leverage (2)

$$\psi(\theta) = \psi(\theta + 2\pi n)$$
$$e^{\pm i\frac{R}{\hbar}\sqrt{2mE}\theta} = e^{\pm i\frac{R}{\hbar}\sqrt{2mE}(\theta+2\pi n)} = e^{\pm i\frac{R}{\hbar}\sqrt{2mE}\theta} e^{\pm i\frac{R}{\hbar}2\pi n\sqrt{2mE}}$$

cancelling the exponential term on both sides, we have

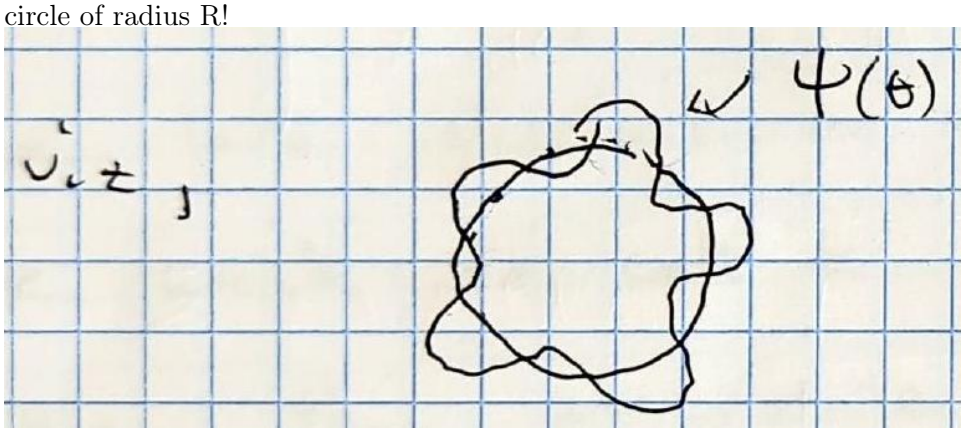
$$1 = e^{\pm i\frac{R}{\hbar}2\pi n\sqrt{2mE}}$$

$$\Rightarrow E = \frac{n^2\hbar^2}{2mR^2}, \forall n \in \mathbb{N}$$

We have officially solved this quantum system (at least spatially, but the time dependent term comes for free! see below)!!

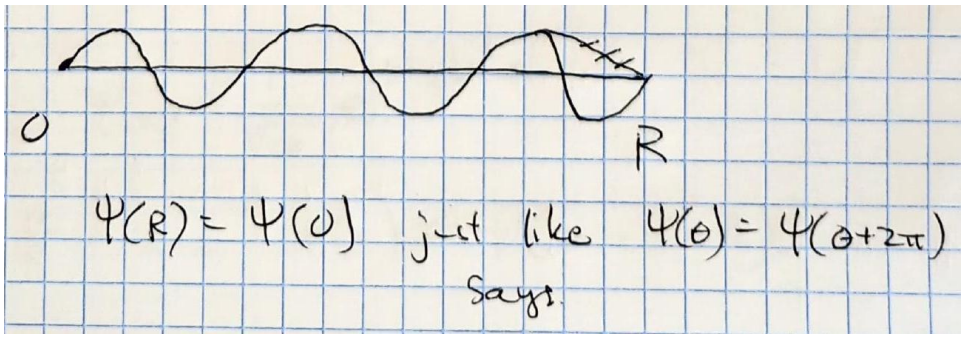
$$\psi_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{\pm in\theta}$$
$$E_n = \frac{n^2\hbar^2}{2mR^2} \quad n \in \mathbb{N}$$
$$\phi_n(t) = e^{\frac{iE_n t}{\hbar}} = e^{i\omega_n t}$$

What does the spatial wavefunction $\psi_n(\theta)$ look like? Well, we know $e^{\pm in\theta}$ are sines and cosines so this is a standing "wave" that's wrapped around a circle of radius R!



Notice that the number of wiggles of the standing wave depends on the energy E_n. The larger E_n the more wiggly the wave. The $\phi_n(t)$ causes these standing wave patterns to translate around the circle in time. The larger the energy, the faster these waves travel around.

Now, imagine if we cut this circle and laid it out flat, what would we expect?



As far as the quantum object is concerned nothing has changed: the boundary conditions are still periodic. (notice that this universe is $2\pi R$ long.

What if we were to make R really really big? We the wavelength of the quantum object would stretch along with R. But this is just fine because, again, as far as the quantum object is concerned this is a perfectly valid state to be in and has a perfectly valid energy.

Now, since we cut the circle and laid it out flat we may as well call the spatial dimension of the $1D$ universe x. So

$$\psi_n(x) = \frac{1}{\sqrt{2\pi}} e^{\pm inx}$$

But now the units dont match! The exponential is no longer dimensionless and i have no clue what e^{meters} means! So that we can knock off the units of x we need a term with dimensions of $\frac{1}{x}$. Let's call it k and since we have cut and stretched our universe (viz, rescaled) it should be $k = \frac{n}{R}$

$$\psi_k(x) = \frac{1}{\sqrt{2\pi}} e^{\pm ikx}$$

where $k = \frac{n}{R}$ and $\psi_k(x) = \psi_k(x + 2\pi R)$. We still require periodic boundary conditions, otherwise what the heck is a wave?

Now, just for fun let's ask "what is the momentum of this wave/state?"

Well, the momentum operator is the generator of spatial translation so

$$\begin{aligned} \hat{p}\psi_k(x) &= -i\hbar \frac{d\psi_k(x)}{dx} \\ &= -i\hbar(ik)\psi_k(x) = \hbar k \psi_k(x) \end{aligned}$$

We find that the momentum is $\hbar k$ and also that $\psi_k(x)$ is actually a momentum eigenstate!? That's really weird because we didn't build this thing using notions of momentum. Well, k depends on the parameter n that periodicity imposed on our solutions. Therefore, because Schrödinger is a linear differential equation then, a sum of solutions is itself a solution and we could in principle have a BUNCH of different values of k_n some of them even many many times over.

Therefore the most general solution to our problem can be written as

$$\psi(x) = \sum_{k=-\infty}^{\infty} \tilde{\psi}_k e^{ikx}$$

Let's unpack:

1.) We sum over k because theres really no difference between k and n

2.) We sum from $-\infty$ to ∞ since our solution was $\pm ikx$. this takes care of both directions

3.) $\tilde{\psi}_k$ multiplies each term of the sum because, as previously mention we could have 2356 copies of a particular value of n or k. This amounts to an amplitude or magnitude for each possible wave k

Now, this is a quantum system so the probability of finding $k = 5$ is given by $|\tilde{\psi}_5|^2$. That's just how quantum wavefunctions work!

Now for some real magic. Let $R \to \infty$. Remember, we had just made it really big but never completely took the limit.

if $R \to \infty$ then $k = \frac{n}{R}$ becomes a $\mathbb{R}$ number rather than just a rational number $\mathbb{Q}$

Now instead of having a bunch of discrete values of k we have a continuum! Therefore our sums turn into integrals and we have

$$\psi(x) = \sum_k \tilde{\psi}_k e^{ik_n x}$$
$$\to \psi(x) = \int_{-\infty}^{\infty} \widetilde{\psi}(k) e^{ikx} dk$$

We have a continuum of k states in an infinite universe. This seems really mathematically shady and indeed it absolutely is. The best I can offer at this level is that it works and everything we did is a matter for mathematicians rather than physicists. Math folks will go on about "distributions", "generalized functions" and "Schwarz space" and so on. You can always look under the hood if you're interested.

Okay, so now we can make sense of continuous k and therefore continuous values of momentum $p = \hbar k$. For example, what is the probability we measure our general state to have momentum p equal to π?

$$\tilde{\psi}^*(k = \frac{\pi}{\hbar})\tilde{\psi}(k = \frac{\pi}{\hbar})$$

However, this $\psi(x)$ is the general SPATIAL wave function of a free particle. How did these momentum probability amplitudes sneak their way in? Don't we ask "what is the probability of the object being at $x = x_0$ at time t?" Well, we simply take

$$\psi^*(x_0)\psi(x_0)$$

and everything works itself out. By all means you should write out the FULL product with all the k's and all!

There seems to be a DEEP relationship between momentum, wave number (k) or momentum, and space.

viz, $\hat{p} \equiv$ generator of spatial translation $= -i\hbar\frac{d}{dx}$ in the position basis of our Hilbert space $\mathcal{H}$

$p = \hbar k \equiv$ eigenvalue of momentum operator.

The connection is that any $\psi(x)$ can be written as a linear combination of momentum basis states! Our abstract quantum state is some vector in Hilbert space. i.e. $|\psi\rangle \in \mathcal{H}$. and $|\psi\rangle$ can be written in terms of the basis $|k\rangle$. Just to visualize whats going on i wrote $|\psi\rangle$ as a sum over discrete k states and then as the proper continuous version over k

$$|\psi\rangle = \alpha_0|k=0\rangle + \alpha_1|k=2\rangle + \dots.$$
$$|\psi\rangle\mathbb{1} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|\psi\rangle|k\rangle\langle k|dk$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\langle k|\psi\rangle|k\rangle dk$$

Where $\langle k|\psi\rangle = \widetilde{\psi}(k)$. That's what $\widetilde{\psi}(k)$ secretly was all along! Now we can project $|\psi\rangle$ onto the $|x\rangle$ basis. We form $\langle x|\psi\rangle$

$$\langle x|\psi\rangle = \psi(x) = \langle x|(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\langle k|\psi\rangle|k\rangle dk)$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\langle k|\psi\rangle\langle x|k\rangle dk$$

But $\langle x|k\rangle = e^{ikx}$ so that

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\langle k|\psi\rangle e^{ikx}dk$$

where $\langle k|\psi\rangle$ is defined to be what we mean by $\widetilde{\psi}(k)$

Question: if we know $\psi(x)$ can we extract $\widetilde{\psi}(k)$?
The answer is an amazing and resounding YES!

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int\tilde{\psi}(k)e^{ikx}dk$$

Cleverly, if we integrate both sides against $e^{-ik'x}dx$ then a beautiful cancellation occurs

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi(x)e^{-ik'x}dx = \frac{1}{\sqrt{2\pi}}\int\tilde{\psi}(k)\int dx e^{i(k-k')x}dk$$

since $\tilde{\psi}(k)$ doesnt depend on x we can do the x-integral first.
Try and prove that:

$$\int e^{i(k-k')x}\,dx = 0 \quad \text{if} \quad k \neq k'$$

and

$$= 2\pi \quad \text{if} \quad k = k'$$

this sounds exactly like 2π times a Dirac delta function $\delta(k-k')$

$$\frac{1}{\sqrt{2\pi}} \int \tilde{\psi}(k) 2\pi \delta(k-k') dk$$

$$= \frac{1}{\sqrt{2\pi}} 2\pi \widetilde{\psi}(k') = \sqrt{2\pi}\widetilde{\psi}(k)$$

$$\widetilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int \psi(x) e^{-ikx} dx$$

given $\psi(k)$ we get $\psi(x)$ and vice versa for free!

This technology of transforming $\psi(x) \rightarrow \widetilde{\psi}(k)$ is called "The Fourier Transform" and is MUCH more general than our quantum object in a $1D$ universe. Generally, for any periodic function we can find a unique way to expand that periodic function as a linear combination of pure plane waves. A plane wave is e^{ikx} where k is the "wavenumber" and is the spatial periodicity of the wave. For example, the wavelength λ of a wave is related to wavenumber by $\lambda = \frac{2\pi}{k}$

What we have done is

$$\widetilde{\psi}(k) = \mathcal{F}[\psi(x)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x) e^{-ikx} dx$$

where $\mathcal{F}$ eats functions and spits out new functions. It is a specific kind of transformation. x and k are called "conjugate" variables. they are related by $\mathcal{F}$. Time and frequency are also conjugate:

$$\widetilde{f}(\omega) = \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(t) e^{-i\omega t} dt$$

Example,
Suppose we know $\psi(x) = \delta(x)$
This says the position of our quantum object is definitely at exactly $x = 0$ with 100% certainty. What is the momentum function $\widetilde{\psi}(k)$.

$$\widetilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(x-0) e^{-ikx} dx$$
$$= \frac{1}{\sqrt{2\pi}} e^{-ik0} = \frac{1}{\sqrt{2\pi}}$$

This says that if x is precisely know with 100% certainty then $\widetilde{\psi}(k)$ is constant

this $\widetilde{\psi}(k) = \frac{1}{\sqrt{2\pi}}$ means there is a uniform probability for finding ANY k! This means knowing position completely ruins any possible knowledge of momentum. it could have $k = -5$ just as easily as it could be $k = 87234892$

If we know x exactly we know NOTHING about k

$$\text{prob}(k=0) = \text{prob}\left(k = 10^{40}\right) = \ldots = \frac{1}{2\pi}$$

This is called "Heisenberg uncertainty" and results from an application of Fourier transformation!. it is also DEEPLY related to the commutator relations $[\hat{x}, \hat{p}] = i\hbar\mathbb{1}$. More to come.

Example, We can even apply Fourier transformations to non-periodic functions.

*Find the Fourier Transform of $f(x) = x^2$ for $0 < x \leq 2$

$f(x) = x^2$ is clearing not perdiodic so if theres any hope of using Fourier's method we will first need to FORCE periodicity.

1. Extend the range to $-\infty < x \leq \infty$
2. Force $f(x) = f(x + 4k)$ where $k \in \mathbb{N}$

Now we have created a periodic function (sketch this!) and furthermore it is
an even function on our extended domain. Now we apply Fourier:

$$\widetilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int f(x) e^{-ikx} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int f(x)(\cos kx - i \sin kx) dx$$

Now, since our $f(x) = x^2$ is even and we're integrating over a symmetric domain $[-\infty, \infty]$ the $f(x) \sin kx$ integrates to zero. i.e. an even function times an odd function is an odd function and the range is symmetric. So we just need the $f(x)\cos(kx)$ term.

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)\cos kx dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2\cos kx dx = \widetilde{f}(k)$$

$$= \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} x^2\cos kx dx$$

This is not a simple integral to compute but can be solved using integration by parts but then the infinite limits are tricky. Try it out, youll find it's infinite unless $k = 0$.

$$= \frac{1}{\sqrt{2\pi}}(\frac{1}{k}\sin kx - \frac{2}{k}(\frac{1}{k^2}\sin kx - \frac{1}{k}x\cos kx)$$

Which must be evaluated at the infinite limits. Alternatively, you can use a table of integrals to find that

$$\mathcal{F}[x^2] = \sqrt{2\pi}\delta''(k-0)$$

As for the second derivative of Dirac we have to appeal to the mathematical theory of "distributions". It turns out that these delta "functions" aren't real functions and when you see one it's usually implied that these "distributions" have to be integrated against true functions.

e.g.

$$\int g(k)\delta''(k-0)dk$$

To handle this kind of expression all one needs to do is integration by parts to transfer the derivatives on $\delta(k)$ to the $g(k)$. This is a well defined expression and procedure but any time you see a $\delta(k)$ by itself you must be a little careful. In physics we tend to be quite cavalier about these "generalized functions" or "distributions".

The lesson here is that we CAN transform non-periodic functions by imposing periodicity within some range, here we forced x^2 to be periodic every $4k$ integer. You can play all sorts of games this way such as making $k \to \infty$ or 0 or whatever else you want.

Now, this is just a demonstrative example, in QM we almost always deal with waves and wavepackets which have a definite periodicity to them. In particular our particle on a circle was defined to be periodic and we let the universe's size go to infinity. this says strictly that we want $\psi(-\infty) = \psi(\infty)$ which is always the case UNLESS youre studying some esoteric topology of the universe which people do and it very well could be possible the universe is non-trivially "twisted" such that $\psi(-\infty) = -\psi(\infty)$.

This would have all sorts of subtle and considerable consequences and I personally have no clue what those consequences would be. The point is, folks study such things usually in the context of quantum field theory where topology is massively important. There are structures such as "cosmic strings", "domain walls", "monopoles", "hedgehogs", and many more exotic quantum objects. In fact, we thoroughly EXPECT such things to be scattered around the universe but we don't see them! The theory of "cosmic inflation" explains why our universe isn't littered with such topological objects (among other things)

Quantum Action

The two slit experiment

Recall than in QM there is a special operator $\hat{H}$ that generates time translations

$$\hat{H}|\psi\rangle = i\hbar\frac{\partial}{\partial t}|\psi\rangle$$

This is a differential equation that is formally solved as

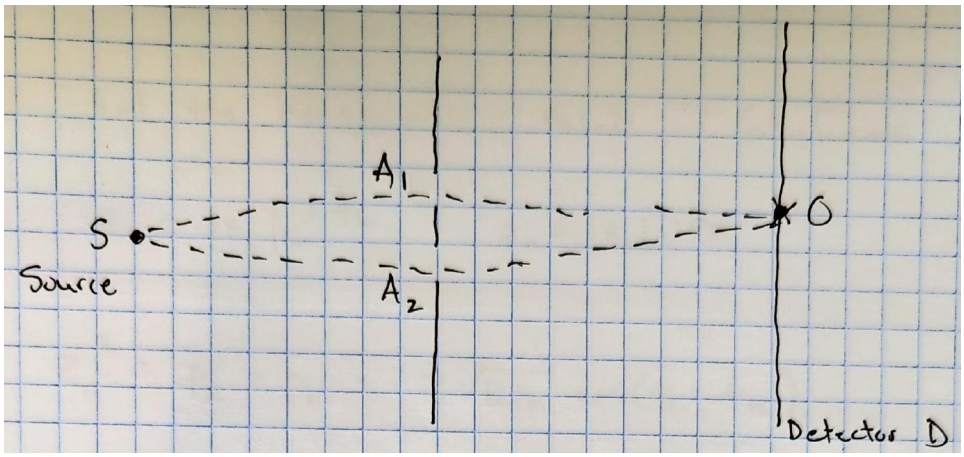
$$|\psi(t)\rangle = e^{\frac{i\hat{H}(t-t_0)}{\hbar}}\,|\psi\,(t_0)\rangle$$

we say

$$U\,(t;t_0) = e^{\frac{i\hat{H}(t-t_0)}{\hbar}}$$

is the "propagator".

Suppose we fire light at a barrier with two slits close together which transmit light. After the slits we record the light with a photographic plate or CCD



We label the slits A_1 and A_2

We ask" what is the amplitude for light at S to propagate to position O

Call the state at position $S\,|q_i\rangle$ and at $O\,|q_f\rangle$. So the quantum object (light or whatever) can travel through A_1 or A_2 the amplitudes for each path add by linear superposition $\therefore$ (amplitude to detect at O).

$$\mathcal{A}\,(S \to A_1 \to O) + \mathcal{A}\,(S \to A_2 \to O)$$

or

$$\mathcal{A}(S \to O) = \sum_i \mathcal{A}\,(S \to A_i \to O)$$

Or if these amplitudes are general complex numbers

$$z_0^2 = |z_L + z_R|^2 = z_L^2 + z_R^2 + z_L^* z_R + z_R^* z_L$$

but z_L^2 = prob (to take path L)

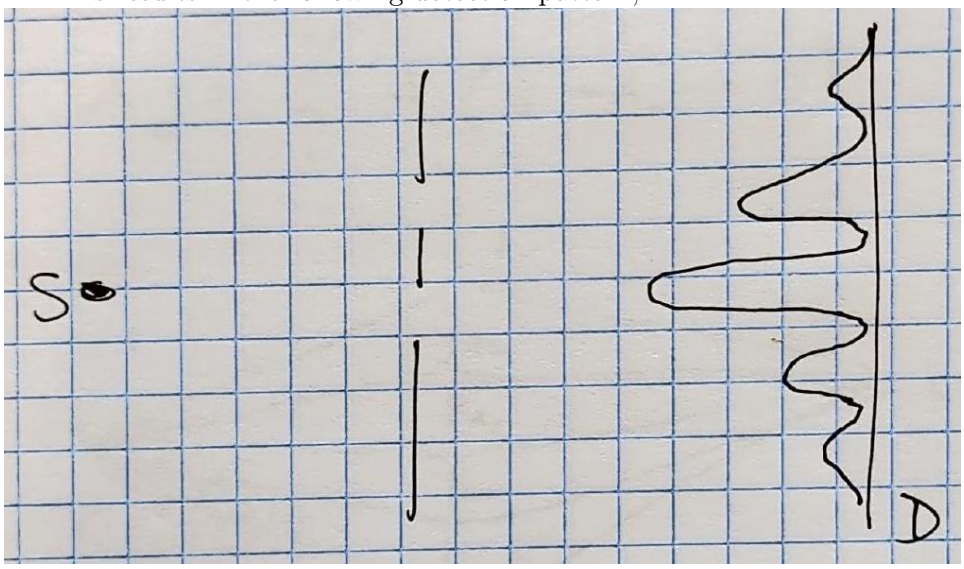
z_R^2 = prob (to take path R)

But the cross-terms result in interference

i.e.

$$z_L = |z_L|\, e^{i\phi_L} \quad z_R = |z_R|\, e^{i\phi_R}$$
$$z_0^2 = p_R + p_L + 2\sqrt{p_L p_R}\cos\left(\phi_L - \phi_R\right)$$

This results in the following detection pattern,



Rather than a simple sum of each path separately

The cross-terms result in interference EVEN if we send in an only 1 object every 3 years!

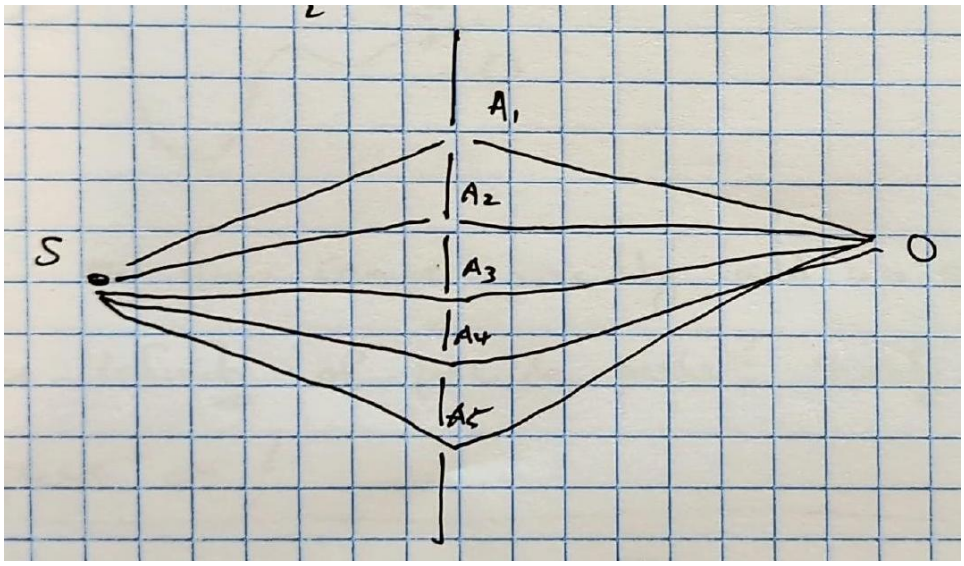
Okay, so $z_0 = z_L + z_R$

or

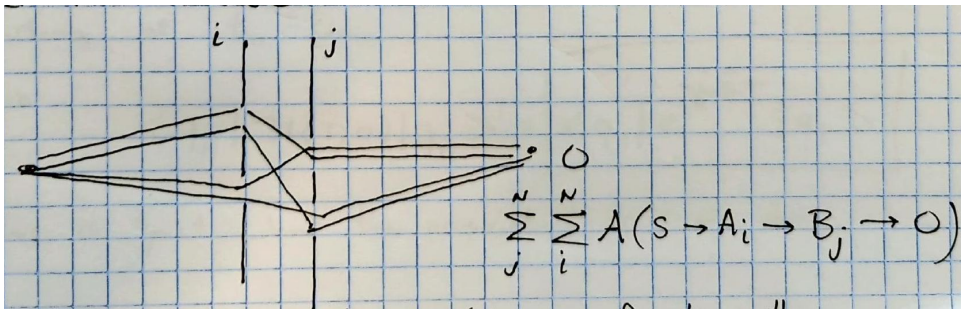
$$\mathcal{A}_O = \sum_i^2 \mathcal{A}\left(S \to A_i \to O\right)$$

What if we add holes, i.e.. 3-slits, 4-slits etc? Well, we simply add up all the probability amplitudes, all N of them.

$$\mathcal{A}_O = \sum_i^N \mathcal{A}\left(S \to A_i \to O\right)$$



What if we have two screens with 2 holes?



We simply add the amplitudes of all possible paths!

$$\mathcal{A}_O = \sum_i^2 \sum_j^2 \mathcal{A}\left(S \to A_i \to B_j \to O\right)$$

What about if there are N holes in each?

$$\mathcal{A}_O = \sum_i^N \sum_j^N \mathcal{A}\left(S \to A_i \to B_j \to O\right)$$

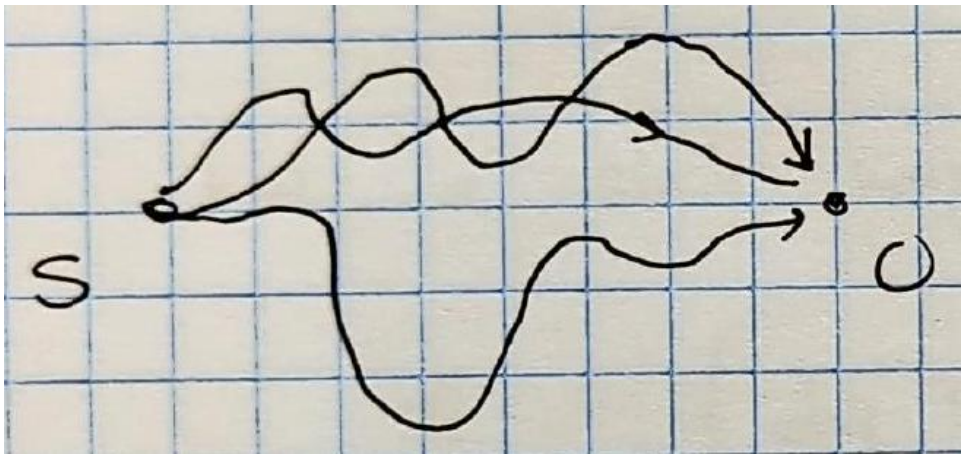
What if we have M screens with N holes each? Well, we would have M separate sums

$$\mathcal{A}_O = \sum_a^N \cdots \sum_m^N \mathcal{A}\left(S \to A_a \to \cdots \to B_m \to O\right)$$

Now, finally suppose we let the number of holes AND the number of screens go to infinity? Well, we would have an infinity of sums over an infinity of holes. What could that even mean? Is this a countable infinity or an uncountable infinity?!

Physically this is the situation of no obstacles at ALL!

$$A(S \to 0) = \sum_{\text{paths}} A_{path}(S \to O)$$



Yikes! how can we possibly add up all the amplitudes of an infinity of possible paths! Surely this is impossible. We press on! Let's try and view this from another point of view;

Using the propagator we can find the amplitude of state $|q_i\rangle$ to evolve into state $|q_f\rangle$ in some amount of time T starting at $t_0 = 0$

$$\langle q_f|U(t;0)|q_i\rangle = \langle q_f|e^{-\frac{i}{\hbar}\hat{H}T}|q_i\rangle$$

This is just what the propagator $U(T;0)$ does. **For now let us relax our notation and set $\hbar = 1$ for notational clarity, we will put it back in its proper place at the very end**.

Dirac's trick: lets chop T into little segments and analyze those

$$\delta t = T/N$$

$$\langle q_f|e^{-i\hat{H}T}|q_i\rangle$$

$$= \langle q_f|e^{-iH\delta t}e^{-iH\delta t}\cdot\cdots\cdot e^{-iH\delta t}|q_i\rangle$$

Next, we are always free to multiply by the identity $\mathbb{1}$ anywhere or everywhere. A slick way of writing $\mathbb{1}$ in quantum theory is

$$\mathbb{1} = \int dq_j\, |q_j\rangle\langle q_j|$$

We say "we take a complete basis of position states" and label them by $j = 1, 2, \cdots$.

Now we put this $\mathbb{1}$ in between all of the $e^{-iH\delta t}$ operators as follows:

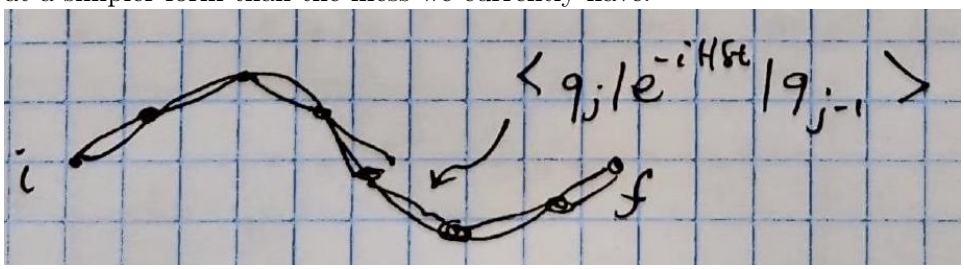
$$\int dq_1 \int dq_2 \ldots \int dq_{N-1}\langle q_f|e^{-iH\delta t}|q_{N-1}\rangle\langle q_{N-1}|e^{-iH\delta t}|q_{N-2}\rangle\langle q_{N-2}|e^{-iH\delta t}\cdots|q_1\rangle\langle q_1|e^{-iH\delta t}|q_i\rangle$$

Or, more compactly,

$$\prod_j^{N-1}\int dq_j\langle q_f|e^{iH\delta t}|\ldots q_j\rangle\langle q_j|\ldots|q_i\rangle$$

This tells us we have a bunch of little paths all stitched together. Each little path amounts to a little probability amplitude. If we can figureout a nicer way to write one of these little amplitudes then perhaps we could arrive

at a simpler form than the mess we currently have!



Okay, one little segment of the path is

$$\langle q_{j+1}| \, e^{-i\hat{H}\delta t} \, |q_j\rangle$$

Now, consider $\hat{H}$ for a freely propagating particle then

$$\hat{H} = \frac{\hat{p}^2}{2m}$$

Since there is no potential.

If we had momentum basis states then we could use

$$\hat{H}|p\rangle = \frac{p^2}{2m}|p\rangle$$

where $\frac{p^2}{2m}$ is a real number.

But wait! we can always use $\mathbb{1}$ for our heart's desires! We desire momentum states! Let there be momentum states!

$$\mathbb{1} = \frac{1}{2\pi}\int dp|p\rangle\langle p|$$

so, $\langle q_{j+1}|e^{-i\hat{H}\delta t}|q_j\rangle =$

$$\frac{1}{2\pi}\int dp|p\rangle\langle p|\langle q_{j+1}|e^{-i\hat{H}\delta t}|q_j\rangle =$$

$$= \frac{1}{2\pi}\int dp\langle q_{j+1}|e^{-i\hat{H}\delta t}|p\rangle\langle p|q_j\rangle$$

$$= \frac{1}{2\pi}\int dp\langle q_{j+1}|e^{-i\frac{p^2}{2m}\delta t}|p\rangle\langle p|q_j\rangle$$

$$= \frac{1}{2\pi}\int dp e^{-i\frac{p^2}{2m}\delta t}\langle q_{j+1}|p\rangle\langle p|q_j\rangle$$

and $\langle p|q_j\rangle = e^{ipq_j}$ and $\langle q_{j+1}|p\rangle = e^{-ipq_{j+1}}$

so that the entire path piece becomes

$$= \int \frac{dp}{2\pi} e^{-i\delta t \frac{p^2}{2m}} e^{ip(q_{j+1}-q_j)}$$

Now, this may look really complicated but if all we care about is integrating over p then this expression amounts to astandard Gaussian integral.

$$= \int \frac{dp}{2\pi} e^{-i\delta t \frac{p^2}{2m}} e^{ip(q_{j+1}-q_j)} = \sqrt{\frac{-2\pi i m}{\delta t}} e^{\frac{im}{2\delta t}(q_{j+1}-q_j)}$$

So each of these mini-paths are Gaussian and equal the above expression and then we simply take the product of all the paths

$$\langle q_f | e^{-i\hat{H}T} | q_i \rangle = \left(\frac{-2\pi i m}{\delta t} \right)^{\frac{N}{2}} \prod_j^{N-1} \int dq_j e^{i\frac{\delta t m}{2} \sum_j^{N-1} \left[\frac{q_{j+1}-q_j}{\delta t} \right]^2}$$

Surely this is the most complicated formula you've ever come across! Observations

1. On the right hand side there are absolutely NO any quantum operators!! ZILCH!
2. Each integral over q is a Gaussian and we generally know how to do Gaussians so maybe there's an analytic solution in there somewhere??
3. We only have a finite number of integrals to do (N-1 of them) so we could, at the very least, program a computer to churn these out.

Now, let's be bonkers for a moment and take $N \to \infty$. We were approximating paths in terms of little segments so we hope as $N \to \infty$ we will reproduce the actual infinitely many continuous paths.

$$\langle q_f | e^{-i\hat{H}T} | q_i \rangle = \lim_{N\to\infty} \left(\frac{-2\pi i m}{\delta t} \right)^{\frac{N}{2}} \prod_j^{N-1} \int dq_j e^{i\frac{\delta t m}{2} \sum_j^{N-1} \left[\frac{q_{j+1}-q_j}{\delta t} \right]^2}$$

- Now Notation

we call $\lim_{N\to\infty} \left(\frac{-2\pi i m}{\delta t} \right)^{\frac{N}{2}} \prod_j^{N-1} \int dq_j = \int \mathscr{D}q(t)$
then,

$$\langle q_f|e^{-i\hat{H}T}|q_i\rangle = \int \mathscr{D}q(t)e^{i\int dt\frac{1}{2}m\dot{q}}$$

is called the Feynman Path Integral. This is notation: it should always be kept in mind that in real life we have the definition of $\int \mathscr{D}q(t)$ with the big products and integrals. Notice that when we take the limit as $N \to \infty$ the terms in the exponent become our standard expression for kinetic energy! In fact, if we had also considered a potential energy then the term being integrated over time would be the Langrangian $\mathcal{L}$!! This is mysterious indeed! There was zero expectation that the Classical Action would emerge from such a quantum thought experiment.

Casually you may consider this as "integrating over functions or paths". There are tricks to evaluating these things and they can be delightful to work with but they are absolutely mathematically ill-defined!

Notice, importantly, there are NO Quantum variables AT ALL in this Path integral formalism! Quantum mechanics can be done via classical action action methods.

In general

$$\langle q_f|e^{-i\hat{H}T/\hbar}|q_i\rangle = \int \mathscr{D}q(t)e^{\frac{iS}{\hbar}}$$

Recall in classical land, we usually want to extremize the action. Look at $e^{\frac{iS}{\hbar}}$. If the action is large for some path then
$e^{\frac{iS}{\hbar}}$ will oscillate rapidly.

va $e^{\frac{iS}{\hbar}} = \cos(\frac{S}{\hbar}) + i\sin(\frac{S}{\hbar})$

When we integrate over the rapidly oscillating terms the integral averages to a tiny number or even zero

Now, if S is near its classical value then these paths will contribute a lot.

The proper method for analyzing integrals such as this is called "the method of steepest decent". We expand our action around the classical path called S_C and then keep terms of order $\delta^2 S$ in the Taylor expansion. Since S_C is just a number it pulls out of the integral and whats left is generally a Gaussian due to $\delta^2 S$.

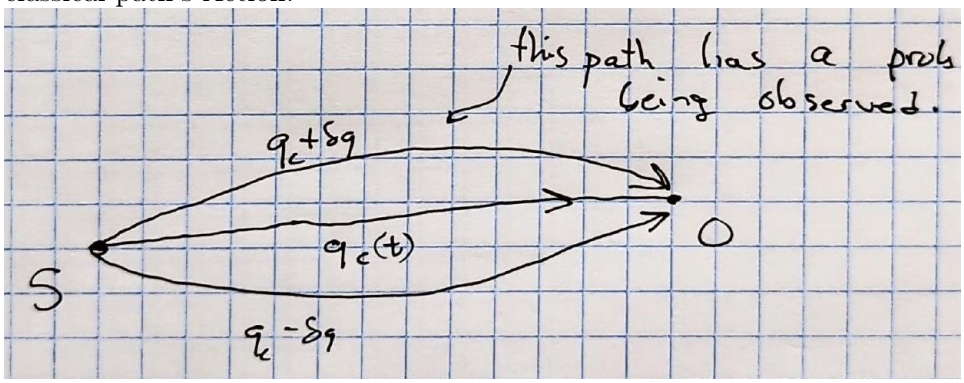
The everyday classical path by far contributes the most amplitude. This says we should usually observe the path of our quantum object to be close to the classical expected path.
However, there will be contributions from paths that are close to but not equal to the classical path from $\delta^2 S$.

e.g

$$\frac{d}{dt}\frac{\partial y}{\partial \dot{q}} - \frac{\partial y}{\partial g} = 0 \Rightarrow q_c(t)$$

$q_c(t)$ is the classical path.

But $q_c(t) + \delta q$ also contributes to the amplitudes which means there's a probability to observe a wacky path so long as its Action is close to the classical path's Action!



This formulation says "consider all possible paths and weight them by $e^{iS/h}$

Then we add up (integrate) all the contribunions in order to find the amplitude for an object to go from $S \to O$
the probability of arriving at O is then

$$\left[\int \mathscr{D}q(t)e^{iS/\hbar}\right]^2 = \text{prob}(S \to O)$$

Now, we wont be making use of this alternative formulation of quantum theory but you should be aware of it. With all this abstract business of Hilbert space and operators and such out there is a completely alternative approach that has no states or operators anywhere in its computational machinery.

Additionally, these Path Integrals are well suited for applications in stochastic calculus such as finance and economics/sociophysics.

Composite Systems and Total Angular Momentum

So far we have learned how to hand single quantum objects yet the questions remains: "How do we handle systems of MANY quantum objects?"'

Additionally, we have learned about two types of angular momentum; spin and orbital. The question emerges, "is total angular momentum simply

the sum of $\hat{L}$ and $\hat{S}$?" What about the states? Do they simply add vectorially or are they something else entirely?

In order to address questions such as this we must postulate a new structure of our theory. This postulate is well motivated mathematically and physically it corresponds to everything we have experimentally measured. It's a structure we sprinkle into our theory.

We say, "if you have two or more quantum systems and you wish to describe them as a single quantum system then you are to form the vector space tensor product of the component systems which represent the composite system."

That is to say, given two Hilbert spaces, say, $\mathcal{H}_A$ and $\mathcal{H}_B$, we form the composite system as

$$\mathcal{H}_{\mathcal{T}} = \mathcal{H}_A \otimes \mathcal{H}_B$$

Let's unpack this abstract expression using our simplest system we know of: qubits.

Suppose I have two 2-level (qubit) systems A and B with the following basis states:

$$\begin{matrix} |1\rangle_A & |1\rangle_B \\ |0\rangle_A & |0\rangle_B \end{matrix}$$

System A has a Hilbert space $\mathcal{H}_A$ and system B has a Hilbert space $\mathcal{H}_B$

We have previously solved a single qubit system and found that the most general state is given by

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

Where two complex numbers α and β fully characterize our single qubit system. our Hilbert space for a single qubit is just $\mathbb{C}^2$.

For two identical qubit systems brought together to form a composite system we say the total Hilbert space

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$$

This says we require four complex numbers to describe a single state of our AB system.

Note that the 2-dimensional space of complex numbers $\mathbb{C}^2 \sim \mathbb{R}^4$

We might anticipate that both A and B can be combined into a basis comprising the "vectors $\in \mathcal{H}_{AB}$"

$$|0\rangle_A|0\rangle_B, \quad |0\rangle_A|1\rangle_B, \quad |1\rangle_A|0\rangle_B, \quad |1\rangle_A|1\rangle_B$$

Such that a "joint" state would be

$$|\psi\rangle = \alpha|0\rangle|0\rangle + \beta|0\rangle|1\rangle + \gamma|1\rangle|0\rangle + \delta|1\rangle|1\rangle$$

This says that there's a state where both qubits are in state $|0\rangle$ a state where the first is in $|0\rangle$ and the second is in $|1\rangle$, a state where vice versa, and a state where both are in $|1\rangle$. Seems reasonable.

Now, there's no condition here that these qubits need interact in any way whatsoever. They could be in opposite ends of the universe. If they have never interacted with each other then there will not be any correlations between them and so measuring A would not yield ANY information about the state of B but mathematically we're allowed to describe them as a composite system but clearly the most interesting situations occur when there is some correlation or interaction between the systems.

At any rate, here we have 4 complex numbers that completely characterize this composite state as expected.

We say $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$

AXIOM: The total Hilbert space of a composite system is the tensor product of the individual systems. e.g. it could be a collection of electrons, or atoms, or humans (if you can describe them quantum mechanically!). We say,

$$\mathcal{H}_{System} = \bigotimes_{i=1}^{N} \mathcal{H}_i$$

This is a postulate of QM that we can do this and nature seems to agree.

Suppose I have two states of two systems

$$|\psi\rangle = \alpha|\alpha\rangle + \beta|\beta\rangle$$

$$|\phi\rangle = \gamma|\gamma\rangle + \delta|\delta\rangle$$

Then,

$$|\psi\rangle \otimes |\phi\rangle = |\psi\phi\rangle = (\alpha|\alpha\rangle + \beta|\beta\rangle)(\gamma|\gamma\rangle + \delta|\delta\rangle)$$

and we multiply everything through, keeping the order or the vectors intact;

$$= |\psi\phi\rangle = |\xi\rangle = \alpha\gamma|\alpha\gamma\rangle + \beta\gamma|\beta\gamma\rangle + \alpha\delta|\alpha\delta\rangle + \beta\delta|\beta\delta\rangle$$

Combining two states from two systems is relatively straight forward. However, the operators we can perform on composite systems require some care.

If we have two systems we might want to operate on A and do nothing to B

Let $|\xi\rangle \in \mathcal{H}_{AB} = H_A \otimes H_B$ be a joint state and let $\hat{S}_A$ act only on A such that

$$|0\rangle_A \to |1\rangle_A \quad \text{and} \quad |1\rangle_A \to |0\rangle_A$$
$$\text{viz} \quad \hat{S}_A = (|0\rangle\langle 1| + |1\rangle\langle 0|)_A$$

then S_A is a two by two matrix

$$\hat{S}_A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

but $|\xi\rangle$ is a $4D$ system meaning all my operators on $\mathcal{H}_{\mathcal{AB}}$ had better be 4 by 4 matrices! How can we manipulate system A with this operator while leaving B alone?

We can do this by leveraging the "direct product" of matrices.

I want $\hat{S}_A$ to act only on A while I do nothing to B
define

$$\hat{S}_{AB} = \hat{S}_A \otimes \mathbb{I}_B$$

$\mathbb{I}_B$ is the 2×2 matrix identity defined on $\mathcal{H}_B$. It does nothing!

Then, for example,

$$\hat{S}_{AB}|01\rangle = |11\rangle$$

the direct product of matrices is defined by

$$S_A \otimes \mathbb{I} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

This is a nice 4 by 4 matrix that acts only on system A (show this by operating on the AB state given by a column of complex numbers!).

Generally, the direct product of matrices is given by:

$$M \otimes T = \begin{bmatrix} m_{11}T & \cdots & m_{1N}T \\ \vdots & & \\ m_{N1}T & \cdots & m_{NN}T \end{bmatrix}$$

Where T is a full matrix. I.e. each element of M multiplies the entire matrix T.

Now, a single 2-state qubit Hilbert space is

$$|\psi\rangle \in \mathcal{H} = \mathbb{C}^2$$

Multiple qubits comprise a Hilbert space of $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$
for n-qubits this is Hilbert space is $\mathcal{H}_n = \mathbb{C}^{2n}$

For $n = 1$ qubit we have 2 basis vectors
For $n = 2$ quits we have 4 basis vectors
For $n = 3$ qubits we have 2^3 basis vectors (e.g $|000\rangle, |010\rangle, \ldots.$)

Generally, we have 2^n possible basis vectors for a system of n qubits

Hilbert Space is insanely big!

If you want to understand just 100 electron spins you'll require
$2^{100} \sim 10^{30}$ states $\Rightarrow 2 \times 10^{30}$ real numbers!

We have no hope at understanding a mol of simple non-interacting electrons! $2^{10^{23}} = ??!$ A mind boggling number of states! In condensed matter physics we work with mol's of electrons all the time and have discovered a LOT. But more remains unknown due to the sheer volume of Hilbert space.

We can write down the full Schrödinger equation for a chunk of matter but we cant solve it analytically without approximations or new mathematics. Maybe an AI can explore such massive Hilbert spaces and find interesting quantum states!

Entanglement

The state $|\psi_A\rangle |\psi_B\rangle = |\psi_{AB}\rangle$ is called a "product state" but not all states are of this form. Consider two systems A and B in the state

$$|\psi\rangle = \frac{|0\rangle_A |0\rangle_B}{\sqrt{2}} + \frac{|1\rangle_A |1\rangle_B}{\sqrt{2}}$$

There is absolutely no way to write this state as an A-only state times a B-only state

$$|\psi\rangle \neq |A\rangle |B\rangle$$

States that are not product states are called entangled states.

Suppose we measure the state of A and find it to be $|1\rangle_A$ then this means we immediately know state B must be $|1\rangle_B$
since $|\psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ then knowing $A \Rightarrow$ knowing B.

They could be on opposite sides of the universe and if we look at A we find out something about the state of B. In fact, measuring A causes B instantly into its corresponding state even though no operation has been performed on B's side of the universe! This isn't necessarily weird in and of itself, it just means the two systems were, at some point in time, locally correlated. For example, you and I could get together and share two coins, one heads and the other tails and not know which we have. The we could go to opposite sides of the Earth and if I see mine is heads then I instantly know yours is tails. Perhaps the entangled state was produced when A and B were together during the big bang and then flew apart billions of light years without their state being disturbed at all. Nothing weird about that.

What can be weird about the entangled system occurs when we decide to measure different properties of the entangled system: for example, choosing at the last moment to measure momentum or position. This kind of set up does lead to the shocking conclusion that our universe violates "local realism". This is the idea that nothing travels faster than light AND objects have definite properties when we aren't measuring them. We know, verifiably the universe violates local realism. It could be that locality is violated and realism is kept, or vice versa, or even both. Nobody knows. More on this later, for now let's further understand composite systems.

Recall our work on the representation theory of $SU(2)$ - the generalized rotations.

We have previously found the generators of $SU(2)$ to be given by the Pauli matrices. $SU(2)$ is given by special unitary matrices. i.e.

$$\mathscr{U}^\dagger \mathscr{U} = 1$$
$$\det \mathscr{U} = 1$$

Its Lie algebra is

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

or, equivalently

$$\left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2}\right] = i\varepsilon_{ijk}\frac{\sigma_k}{2}$$

This is tantalizingly similar to what we found for $SO(3)$ in $SO(3)$

$$[L_i, L_j] = i\varepsilon_{ijk}L_k$$

is here

$$[S_i, S_j] = i\varepsilon_{ijk}S_k$$

where $S_i = \frac{\sigma_i}{2}$

If we go through the same procedure as we did for $SO(3)$ we find all the representations for this $SU(2)$ Lie Algebra. We call these representations "spin". The states are the eigenvectors of the generators of $SU(2)$ which are generally $(2s+1) \times (2s+1)$ matrices for $s = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Spin-$s$ can be half integer valued. We say such a system is an object that has "spin-s"

Recall $SO(3)$ were labeled by integers $l = 0, 1, \ldots$ and $SU(2)$ subsumes the $SO(3)$ integer representations.

For example,

spin-0 $\Rightarrow$ one basis state $|0\rangle$

spin-1/2 $\Rightarrow$ two basis states $|-1/2\rangle, |1/2\rangle$ otherwise known as $|\uparrow\rangle, |\downarrow\rangle$

spin-1 $\Rightarrow$ 3 basis states $|-1\rangle, |0\rangle, |1\rangle$

spin-3/2 $\Rightarrow$ 4 basis states $|-3/2\rangle, |-1/2\rangle, |1/2\rangle, |3/2\rangle$

and so on.

A quantum system can have two types of angular momentum given by "Spin" and "Orbital" angular momentum.

Spin comes from $\mathfrak{su}(2)$ elements and orbital angular momenta from $\mathfrak{so}(3)$ elements.

For example, an electron in carbon can have $\ell = 1$ and $s = \frac{1}{2}$

This is known as a p-orbital in chemistry and the electron can be in an up or down configuration inside that orbital.

We notice that spin-1 overlaps with $\ell = 1$ in $SO(3)$. In fact all of the integer spin states look exactly like the orbital states in $SO(3)$

This allows us to combine angular momentum into a total angular momentum

$$\vec{J} = \vec{L} + \vec{S}$$

Where the vector symbol means a "vector of operators". For example

$$\vec{L} = (\hat{L}_x, \hat{L}_y, \hat{L}_z)$$

Unfortunately we cant "simply" add such operators together. There's a BIG difference between an $s = \frac{3}{2}$ state and an $\ell = 5$ state (the vector spaces are

entirely different sizes!). Operators can be added but the resulting states are NOT simple sums or products.

For example, we may wish to find the possible spin angular momentum states of two electrons. Well, each can be written as a linear combination of $|\uparrow\rangle$ and $|\downarrow\rangle$ spin-$\frac{1}{2}$ states.

$$\begin{aligned}|\psi_1\rangle|\,\psi_2\rangle &= (\alpha|\uparrow\rangle + \beta|\downarrow\rangle)(\gamma|\uparrow\rangle + \delta|\downarrow\rangle \\ &= \alpha\gamma|\uparrow\uparrow\rangle + \alpha\delta|\uparrow\downarrow\rangle + \beta\gamma|\downarrow\uparrow\rangle + \beta\delta|\downarrow\downarrow\rangle\end{aligned}$$

Okay. this is all fine and straight forward but what if we ask the question "what are the possible angular momentum eigenstates of this pair of spin-1/2 objects?". It's not as simple as saying "well this one is up and that one is down" and calling it a day. The system behaves as a single total object and in quantum physics we cannot necessarily "see" single electrons let alone two single electrons paired together somehow. We can, however, measure the quantum states no matter what their composition is. Dear reader, for a VERY long time it wasn't known that a neutron is composed of many individual sub-nuclear objects called gluons and quarks. Yet we were happily able to write down the wavefunction of neutrons and study their properties.

We need a way to categorize the angular momentum states of something inherently complicated like a pair of electrons or even a neutron.

Now, I cant possibly tell what the angular momentum states of the composite $|\psi_1\rangle\,|\psi_z\rangle = |\xi\rangle$ are just by looking and each component individually! is $|\xi\rangle$ spin $\frac{1}{2}$, spin 1, spin 18? What representation of Spin angular momentum does $|\xi\rangle$ inhabit??

If we are methodical in this (as we will prove below) we will find the following angular momentum states of the Hilbert space of the composite system

$$\left.\begin{array}{rc} |s=1,1\rangle = & |\uparrow\uparrow\rangle \\ |s=1,0\rangle = & \frac{1}{\sqrt{2}}|\uparrow\downarrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\uparrow\rangle \\ |s=1,-1\rangle = & |\downarrow\downarrow\rangle \end{array}\right\} \; Spin1$$

$$\left.\begin{array}{rc} |s=0,0\rangle = & \frac{1}{\sqrt{2}}|\uparrow\downarrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\uparrow\rangle \end{array}\right\} \; Spin0$$

The top three states are Spin-1 and the last state comprises a spin-0 representation. The states of spin angular momentum of our composite system are linear combinations of our original sub-system spin states. Evidently when we combine two spin $\frac{1}{2}$ object we get a spin-1 representation and a spin-0 representation. Yes, you should be enormously confused at this point

but bear with me. These particular states are eigenvectors of the total spin operator (in the z direction) acting on the total Hilbert space, i.e.

$$\hat{S}_z^{\text{total}} |\xi\rangle \in \mathcal{H}_{AB}$$

We write abstractly

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$$

This says two spin-$\frac{1}{2}$ representations combined to produce a spin-1 and a spin-0.

Spin-1 has three eigenstates (written above) and Spin-0 has only one.

This

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$$

is sometimes written in terms of the dimension of the irreducible representation (irrep) i.e. $dim = d_j = 2j + 1$

$$d_{\frac{1}{2}} = 2 \quad d_1 = 3 \quad d_0 = 1$$

$$2 \otimes 2 = 3 \oplus 1$$

It's super confusing and both formats have their own merits. I prefer $\frac{1}{2} \otimes \frac{1}{2} = 1 \otimes 0$ because I think of electrons as spin-$\frac{1}{2}$ objects rather than the dimensions of their Hilbert spaces. We will use this standard unless otherwise noted.

In full generality we will find how to determine the angular momentum states of an arbitrary product of states from any representation we like as

$$j \otimes j' = (j + j) \oplus \left(j + j' - 1\right) \oplus \cdots \left(\left|j - j'\right| + 1\right) \oplus \left(|j - j'|\right)$$

e.g. For example, if wee have a total system given by the product of a spin $\frac{1}{2}$ combined with spin-2 then

$$\frac{1}{2} \otimes 2 = \frac{5}{2} \oplus \frac{3}{2}$$

We get a spin-$\frac{5}{2}$ and spin-$\frac{3}{2}$ set of spin angular momentum states.

or in terms of dimensions

$$2 \otimes 5 = 6 \oplus 4 \text{ and}$$

Or how about a spin-1 with spin-1

$$1 \otimes 1 = 2 \oplus 1 \oplus 0$$

or in terms of the dimensions

$$3 \otimes 3 = 5 \oplus 3 \oplus 1$$

Question: What are the spin states of the Hydrogen nucleus? What about the whole atom?

Neutrons and protons are both spin-$\frac{1}{2}$ objects

Let's suppose the nucleus of Hydrogen has one proton and one neutron. Then we have $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$ states for the hydrogen nucleas. If we measure the angular momentum of the nucleas (say, using MRI) then we will measure $s = 1$ or $s = 0$.

Now, whole atom (electron, proton, and neutron) is

$$\begin{aligned} \frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} &= (1 \oplus 0) \otimes \frac{1}{2} \\ &= 1 \otimes \frac{1}{2} \oplus 0 \otimes \frac{1}{2} \\ &= \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \end{aligned}$$

The whole atom of Hydrogen with one e^-, p, and n acts like spin-$\frac{1}{2}$ or spin-$\frac{3}{2}$ system. Fractional spin systems are called "Fermions" and have unique properties.

If, however, there is no neutron then Hydrogen is

$$H : \frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$$

Integer spin objects are called "Bosons". Bosons and Fermions have wildly different behaviors as we will see later.

Example: Super-Conductivity

The BCS mechanism produces "Cooper pairs". This is a pair of electrons bound to each other via vibrational quanta called phonons.

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0 \Rightarrow \text{ boson.}$$

Okay, now for the gritty details.

Procedure for "adding" angular momenta

The following procedure is known under the misleading name of "adding

angular momenta". It, unfortunately, is not so straightforward as the name would lead us to believe. We shall follow the steps below. If at any point you get all mixed up and confused pick specific values for m_ℓ and ℓ. In this discussion we will not be putting hats on operators.

1. Let $|\ell, m_\ell\rangle$ label a state with spin-ℓ where $m_\ell = -\ell, \ell+1, \ldots, \ell-1, \ell$

For example, for $\ell = 1$ we have three eigenvectors $|-1\rangle, |0\rangle, |1\rangle$ which are labeled by the three values of $m_1 = -1, 0, 1$

All we have done is simply changed notation and so now we write

$$\begin{aligned} |-1\rangle &\equiv |1,-1\rangle \\ |0\rangle &\equiv |1,0\rangle \qquad\qquad \text{here } \ell = 1 : m_\ell = -1, 0, 1 \\ |+1\rangle &\equiv |1,1\rangle \end{aligned}$$

We need to do this because if $\ell = 2$ there are states $|-2\rangle, |-1\rangle, |0\rangle, |1\rangle, |2\rangle$ and we cant let ourselves get mixed up between the $|0\rangle$ state from $\ell = 1$ and the $|0\rangle$ from $\ell = 2$. They VERY different states!! So, to combat this, we merely label all states in terms of angular momentum ℓ as well as m_ℓ

2. Now we wish to take two states from potentially different representations

$$|\ell, m_\ell\rangle \otimes \left|\ell', m_{\ell'}\right\rangle$$

We would love to find $L_z^{\text{total}} = L_z^A \otimes \mathbb{I}_B + \mathbb{I}_A \otimes L_z^B = L_z^T$ where A signifies the first state and B is the second state. This L_z^T will act in turn on both states as

$$\begin{aligned} &L_z^T \left|\ell, m_\ell\right\rangle \otimes \left|\,\ell', m_{\ell'}\right\rangle \\ &= \left(L_z^A \left|\ell, m_\ell\right\rangle\right) \otimes \left|\ell', m'_\ell\right\rangle + \left|\ell, m'_\ell\right\rangle \otimes \left(L_z^B \left|\ell', m_{\ell'}\right\rangle\right) \end{aligned}$$

This is how we "add" angular momenta. We pick the z-component and apply it to each system in turn. Now we know how to aply these operators to their respective states. These are all eigenstates of L_z after-all.

If we were instead using L_x then everything would be MUCH more complicated because $|\ell, m_\ell\rangle$ are ONLY eigenvectors of L_z, L_x have entirely different eigenvectors!

Now,

$$L_z^A \left|\ell, m_\ell\right\rangle = m_\ell \left|\ell, m_\ell\right\rangle$$

and

$$L_z^B|\ell', m_{\ell'}\rangle = m_{\ell'}|\ell', m_{\ell'}\rangle$$

$L_z^{Total}|\ell, m_\ell\rangle \otimes |\ell', m_{\ell'}\rangle = (m_\ell + m_{\ell'})|\ell, m_\ell\rangle \otimes |\ell', m_{\ell'}\rangle$

as we might hope the z-components simply add IF we choose L_z to be diagonal (as is traditionally the case).

For example, take

$$\ell = \frac{1}{2}, \ell' = \frac{1}{2}$$

$$m_\ell = \frac{1}{2}, m_{\ell'} = -\frac{1}{2}$$

Then,

$$L_z^T|\frac{1}{2}, \frac{1}{2}\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle = (\frac{1}{2} - \frac{1}{2})|\frac{1}{2}, \frac{1}{2}\rangle|\frac{1}{2}, -\frac{1}{2}\rangle = 0$$

This is the z-component of the angular momentum of an up and down electron combined together. $L_z^T|\uparrow\rangle|\downarrow\rangle = 0$ as we might expect. They cancel each other out. This is a SPECIFIC case and the two electrons can be in a linear combination of $|\uparrow\rangle$ and $|\downarrow\rangle$

So that, in general we could have 4 possibilities for two spin $\frac{1}{2}$ objects,

$$|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$$

One of these four states has the maximum possible total angular momentum (z-component). Clearly the up-up state has the largest value of L_z^{Total}

$$L_z^{Total}|\uparrow\uparrow\rangle = (\frac{1}{2} + \frac{1}{2})|\uparrow\uparrow\rangle = 1|\uparrow\uparrow\rangle$$

Recall that we previously built the raising and lowering operators of angular momentum as

$$L_+ = L_x + iL_y$$

$$L_- = L_x - iL_y$$

We say that these particular operators move us up and down the ladder of m_ℓ states.

We can apply

$$L_-^{Total} = L_- \otimes \mathbb{I}_B + \mathbb{I}_A \otimes L_-$$

to $|\uparrow\uparrow\rangle$ in order to get the state below this "top" state. Let's see what we get, whatever it is it will have $m_\ell^{total} = 0$ because the up-up state had $m_\ell^{total} = 1$ and L_- decreases m_ℓ^{total}. So, proceeding

$$L_-^T|\uparrow\uparrow\rangle = (L_-|\uparrow\rangle)|\uparrow\rangle + |\uparrow\rangle L_-|\uparrow\rangle$$

$$= |\downarrow\rangle|\uparrow\rangle + |\uparrow\rangle|\downarrow\rangle)$$

Therefore, we have

$$L_-^T|\ell^{total}, m_\ell^{total}\rangle = L_-^\top|\frac{1}{2}+\frac{1}{2}, \frac{1}{2}+\frac{1}{2}\rangle = L_-^T|1,1\rangle$$

$$L_-^T|1,1\rangle = \sqrt{2}|1,0\rangle$$

$$\Rightarrow |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$$

The state $|1,0\rangle$ is the state of TOTAL angular momentum(z-component). It is the eigenstate of L_z^T that has eigenvalue 0. Unfortunately the notation of this subject is atrocious. What we have found is that the TOTAL angular momentum state with eigenvalue 0 is given by the above linear combination of ups and downs of the A and B systems. The TOTAL angular momentum state with eigenvalue = 1 occurs when both subsystems are up-up.

Again, this combination of ups and downs is an $\ell^{total} = 1$, $m_{\ell=1}^{total} = 0$ state of $L_z^{Total} = L_z \otimes \mathbb{I}_B + \mathbb{I}_A \otimes L_z$

It's VERY easy to get lost in the weeds and the best path forward is to simply write everything out by hand several times using various specific examples. We have two more states of the composite system to find.

To recap our method so far: we have taken two spin $\frac{1}{2}$ electrons and found the "highest weight" state $|\uparrow\rangle|\uparrow\rangle$. This is the $\left|\ell^{\text{total}}, m_\ell^{\text{total}}\right\rangle = |1,1\rangle$ state of total angular momentum $L_z^{\text{total}} = L_{zA} \oplus L_{zB}$.

$|\uparrow\uparrow\rangle$ behaves exactly like the $\ell = 1$, $m_{\ell=1} = 1$ angular momentum state of a single spin-1 object (even though it is a composite system).

Next we hit this "highest state" $|11\rangle$ with the "lowering operator" L_- to get

$$L_-^{total}|11\rangle = \sqrt{2}|1,0\rangle$$

That's simply what L_-^{total} does
when we apply L_-^{total} on $|\uparrow\uparrow\rangle$ we got

$$L_-^{total}|\uparrow\uparrow\rangle = (L_-|\uparrow\rangle)\,|\uparrow\rangle + |\uparrow\rangle\,(L_-|\uparrow\rangle)$$

$$= |\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle$$

$$\therefore \quad |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$$

This particular combination of electrons behaves exactly like the $\ell = 1$, $m_\ell = 0$ angular momentum state of a spin-1 object.

Now, to press forward, we hit $|1,0\rangle$ with L_-^{total} again! We keep lowering our states until there's nothing left to lower and we get 0. So, following our representation theory rules

$$\begin{aligned} & L_-^{total}|1,0\rangle = \sqrt{2}|1,-1\rangle \\ & |1,0\rangle = \tfrac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ \therefore \quad & |1,-1\rangle = |\downarrow\downarrow\rangle \end{aligned}$$

We may have expected something different but L_-^{total} hits both spins and they combine to $-\frac{1}{2} - \frac{1}{2} = -1 = m_{\ell=1}^{\text{total}}$. We have found that two spin-$\frac{1}{2}$ systems combine into an $\ell^{total} = 1$ spin-1 representation as

$$\begin{cases} |1,1\rangle = |\uparrow\uparrow\rangle \\ |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |1,-1\rangle = |\downarrow\downarrow\rangle \end{cases}$$

But we're not done! There is one final orthogonal state all by its lonesome (remember our Hilbert space of two objects needs four bases vectors and we've only found three!). You can hit $|1,-1\rangle$ with L_-^T or you can say "I just need an orthonormal vector to these three I've already found". Either way you find

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle)$$

this is an $\ell^{total} = 0$ state. No angular momentum at all. Zilch! Call it $|0,0\rangle$. Interestingly it is anti-symmetric if we switch the electrons.

We knew we had four total states $|\uparrow\uparrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\downarrow\rangle$, and these (above) particular combinations of those states furnish the angular momentum states of the totality of the two electron system. If we were to measure

angular momentum of our two electron system these are the states we would find them in.

This is all abstractly laid out when we write $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$. The 1 is a "triplet" of states (which happen to be symmetric under electron exchange!) and the 0 is an anti-symmetric "singlet".

It's a HUGE algebraic MESS to get all the $\sqrt{2}$'s and various factors correct. This whole technology is called Clebsch-Gordon Decomposition and there are tables and tables that tell you how to combine everything including all the numerical constants involved so that you don't have to repeat this process over and over and over.

At this level all we really need to know is that:

For orbital angular momentum:

$$\ell \otimes \ell' = (\ell + \ell') \oplus (\ell + \ell' - 1) \oplus \ldots \oplus |\ell - \ell'|$$

For Spin angular momentum:

$$s \otimes s' = (s + s') \oplus (s + s' - 1) \oplus \ldots \oplus |s - s'|$$

For TOTAL angular momentum $J = L + S$

$$j \otimes j' = (j + j') \oplus (j + j' - 1) \oplus \ldots \oplus |j - j'|$$

All types of angular momenta follow this same exact pattern of composition.

For example,

$$12 \otimes 10 = 22 \oplus 21 \oplus \ldots\ldots 2$$

and so on.

It's pretty straight forward if we don't need all the Clebsch-Gordon factors and specific sub-states.

Example: an electron in a p orbital will have total angular momentum of

$$j = \ell + s \quad \text{or} \quad j = \ell - s$$

Since the electron can be $\pm\hbar\frac{1}{2}$

For p orbitals $\ell = 1$

so the total angular momentum is

$$1 \pm \frac{1}{2} = \frac{3}{2}, \frac{1}{2}$$

or $1 \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2}$

The $\frac{3}{2}$ has 4 states

The $\frac{1}{2}$ has 2 states which makes six total states.

Recall chemistry class where we learn about the periodic table. We learn the p-orbital has 6 slots for up and down electrons and here we have found $6 = 4 + 2$ total states!

We fill the orbital with up and down electrons as below

$$\underset{p_x}{\uparrow\downarrow}\ \underset{p_y}{\uparrow\downarrow}\ \underset{p_z}{\uparrow\downarrow}$$

These particular states above are NOT angular momentum states. Angular momentum states break apart into two classes: spin-$\frac{3}{2}$ and spin-$\frac{1}{2}$ and that's not what the above diagram represents. These two classes are angular momentum so in your mind you may (falsely) imagine a little spinning electron orbiting the nucleus: these are the possible states of angular momentum of that spinning and orbiting electron.

Deconstructing spinors and electrons in magnetic fields

If you're like me then this spinor business seems pretty unrealistic.

How/why does our typical vector

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

get packaged into the matrix

$$\mathbb{X} = \begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix}$$

When you see a matrix written out the first thing you should ask is "what are the basis vectors and what do they represent?". In this case it's not exactly clear. Let's try and find out.

Our goal is to find two spinors such that their outer product equals our matrix
viz

$$\mathbb{X} = \xi\chi^T = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \begin{bmatrix} -\xi_2 & \xi_1 \end{bmatrix} = \begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix}$$

This expression comes from the definition of the outer product, viz,

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 \\ u_2 v_1 & u_2 v_2 \end{bmatrix}$$

For everything to match up we require

$$\begin{array}{cc} -\xi_1 \xi_2 = z & -\xi_2^2 = x + iy \\ \xi_1^2 = x - iy & \xi_2 \xi_1 = -z \end{array}$$

$$\begin{aligned} \xi_1 \xi_2 &= z \\ \xi_1 &= \sqrt{x - iy} \\ \xi_2 &= \sqrt{x + iy} \\ z &= -\sqrt{x - iy}\sqrt{x + iy} \end{aligned}$$

Our two basis spinors are

$$\vec{\xi} = \begin{bmatrix} \sqrt{x - iy} \\ \sqrt{x + iy} \end{bmatrix} \quad \vec{\chi} = \begin{bmatrix} -\sqrt{x + iy} \\ \sqrt{x - iy} \end{bmatrix}$$

So long as $z = -\sqrt{x + iy}\sqrt{x - iy}$ then this will work. However, we have a problem. This says

$$-z^2 = x^2 + y^2$$

or

$$x^2 + y^2 + z^2 = 0$$

These are called "isotropic" vectors. Verifiably this is NOT our reality. How can we possibly describe arbitrary x, y, z using this whole spinor formalism?

With a bit of frustration, confusion, time, and thought we realize the following fact:

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

Perhaps we can re-write our matrix in terms of things that look like this. viz,

$$\begin{bmatrix} z & x - iy \\ x + iy & -z \end{bmatrix} = z \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (x - iy) \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + (x + iy) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + (-z) \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Now we realize that each matrix can be factored as $\xi_i \xi_j^T$ using

$$\xi_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \xi_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\mathbb{X} = \sqrt{z}\xi_1\xi_1^T\sqrt{z} + \sqrt{x-iy}\xi_1\xi_2^T\sqrt{x-iy}$$
$$+ \sqrt{x+iy}\xi_2\xi_1^T\sqrt{x+iy} + i\sqrt{z}\xi_2\xi_2^T i\sqrt{z}$$

We did it! It's not very pretty but now we have our (x, y, z) written in terms of spinors such that each piece of $\mathbb{X}$, the ξ_1 and ξ_2 transform separately under SU(2) as

$$\mathscr{U}\xi_i = \xi_i'$$
$$\xi_i^\dagger \mathscr{U}^\dagger = \xi_i'^\dagger$$

All of this hard work just to say that there exists a map from 3D space to spinor space such that things rotate properly. The spinors themselves rotate through half angles relative to our typical rotations. These spinors are very much like "the square root of a vector". This breakdown and analysis is very hard to come by in textbooks but I find it's crucial to solidify an understanding and intuition behind spinors in physics.

So, we've learned that in QM objects can have an "intrinsic angular momentum" called spin. For spin s there are $2s + 1$ states of that spin.

A complete set of states is sometimes called a "multiplet". These states all have the same energy unless interactions split them apart. Let us see this splitting in action. An $|\uparrow\rangle$ electron and $|\downarrow\rangle$ electron out and about in empty free space have the exact same energy. However, things change when they encounter a magnetic field! Recall from you course on electricity and magnetism that an electron with orbital angular momentum creates a current which induces a magnetic field. Perhaps this "intrinsic angular momentum" we call spin also makes the electron behave like a little magnet even when it has zero orbital angular momentum (indeed it does!).

For our analysis let us take the electron to be motionless and further we will line everything up with the z-axis (coordinates are our choice after-all). The magnetic dipole moment of an electron due to its spin is given by

$$\vec{\mu} = \mu_z = \gamma\hbar\frac{\vec{\sigma}}{2} = \gamma\vec{S}$$

where γ is a constant that depends on the electron. Our electron here is in some random direction, say, at angle ϕ with respect to the z-axis.

Now, recall the interaction potential energy between a magnetic dipole and a magnetic field. Let's take our magnetic field to also be only in the z direction

$$\vec{B} = \begin{bmatrix} 0 \\ 0 \\ B_z \end{bmatrix}$$

Then our Hamiltonian is

$$\hat{H} = -\vec{\mu} \cdot \vec{B} = -\gamma B_z \hat{S}_z$$

All the information about whatever S_x and S_y could be dont contribute to the potential energy. Only the magnetic field's direction. If our B-field was in some arbitrary direction then, sure, there would be interactions with all the various spin components. But because we line up our axes with the B-field we need not consider other components!

Now, we wish to solve

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

for the generic electron state

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$$

where α and β depend on the angle ϕ our spin vector $\vec{S}$ makes with the z-axis.

We plug in and notice that the only operator is $\hat{S}_z$ and we know how that acts on $|\uparrow\rangle$ and $|\downarrow\rangle$

$$\hat{S}_z|\uparrow\rangle = \frac{\hbar}{2}|\uparrow\rangle$$

$$\hat{S}_z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle$$

So that

$$\hat{H}|\uparrow\rangle = -\gamma B_z \frac{\hbar}{2}|\uparrow\rangle$$

and

$$\hat{H}|\downarrow\rangle = \gamma B_z \frac{\hbar}{2}|\downarrow\rangle$$

We find two distinct energies $E_\pm$ where

$$E_+ = -\gamma B_z \frac{\hbar}{2}$$

$$E_- = \gamma B_z \frac{\hbar}{2}$$

Evidently a down electron has a higher energy than an up electron in a magnetic field. I you have an electron it will try to align itself in the direction of a magnetic field. COOL!

Now, that's only part of the solution. We wish to find the time dependence of the electron state in a magnetic field and for this we need the time-dependent Schrodinger equation:

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

whose solution is

$$|\psi(t)\rangle = U(t;0)|\psi(0)\rangle$$

where

$$U(t;0) = e^{-i\hat{H}t/\hbar}$$

is our propagator.

At $t = 0$ we have $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ and since up and down are energy states we ultimately find

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi\rangle$$

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}(\alpha|\uparrow\rangle + \beta|\downarrow\rangle)$$

$$= \alpha e^{-iE_+t/\hbar}|\uparrow\rangle + \beta e^{-iE_-t/\hbar}|\downarrow\rangle)$$

and plugging in $E_\pm$ we get our solution

$$\alpha e^{i\gamma B_z t/2}|\uparrow\rangle + \beta e^{-i\gamma B_z t/2}|\downarrow\rangle)$$

or finally

$$|\psi(t)\rangle = \alpha e^{i\omega t/2}|\uparrow\rangle + \beta e^{-i\omega t/2}|\downarrow\rangle)$$

where $\omega = \gamma B_z$ is called the Larmor frequency.

Let's unpack; what does this say? This says that the probability amplitudes of each state oscillate as a function of time. when $t = 0$ we reproduce α and β but as time ticks on these complex numbers change by a phase. As far as measurement is concerned we measure no difference in probability with respect to S_z. However, S_x and S_y is another story.

We can compute the average spin angular momenta in each direction and get a picture of what our electron state does in such a situation. We compute

$$\langle \hat{S}_x \rangle = \langle \psi(t) | \hat{S}_x | \psi(t) \rangle$$

$$\langle \hat{S}_y \rangle = \langle \psi(t) | \hat{S}_y | \psi(t) \rangle$$

$$\langle \hat{S}_z \rangle = \langle \psi(t) | \hat{S}_z | \psi(t) \rangle$$

Let us write our α and β in terms of ϕ and to ease our computation we will write $|\psi(t)\rangle$ as a column vector

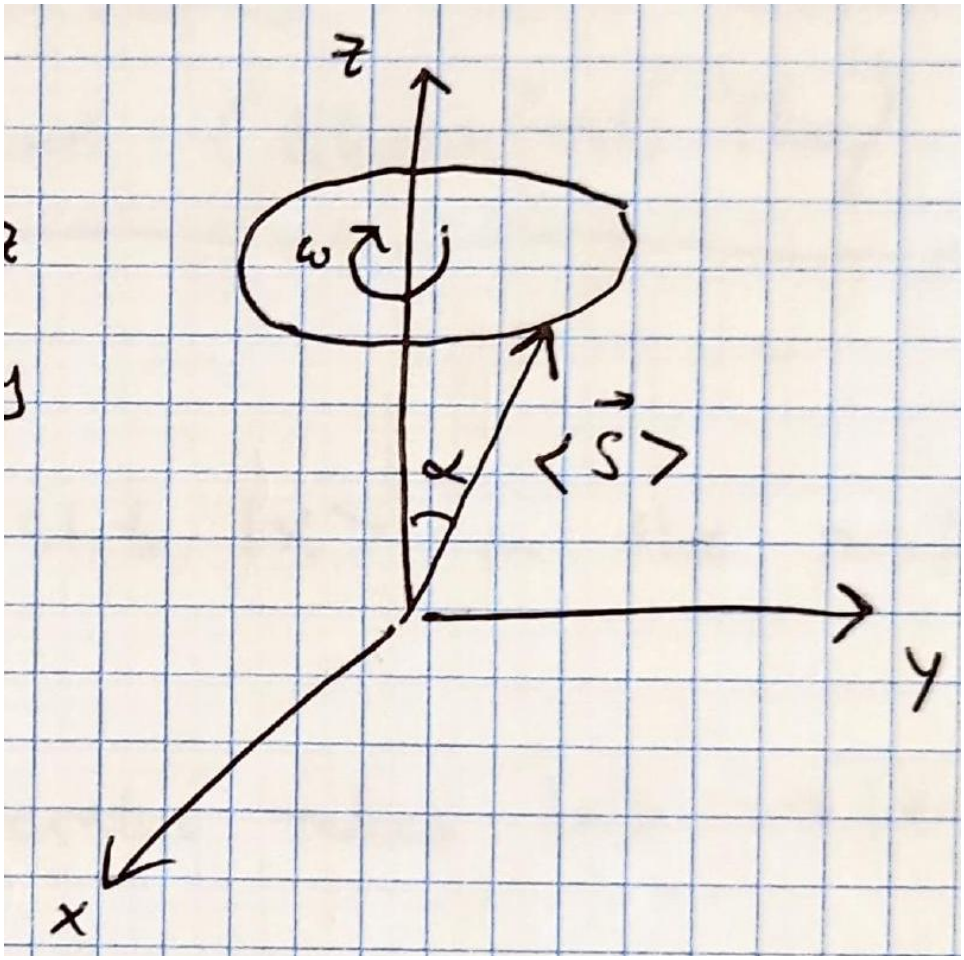
$$|\psi(t)\rangle = \begin{bmatrix} \cos(\frac{\phi}{2}) e^{i\omega t/2} \\ \sin(\frac{\phi}{2}) e^{-i\omega t/2} \end{bmatrix}$$

Now, plugging in the

$$\langle \hat{S}_x \rangle = \frac{\hbar}{2} \sin\phi \cos(\gamma B_z t)$$

$$\langle \hat{S}_y \rangle = -\frac{\hbar}{2} \sin\phi \sin(\gamma B_z t)$$

$$\langle \hat{S}_z \rangle = \frac{\hbar}{2} \cos\alpha$$



Notice that the x and y components of spin are constantly changing as a function of time with frequency γB_z whereas the z component is constant. This is a precession around the z axis. Quantum mechanically the probabilities of the x and y components are constantly changing. At certain times you will measure zero chance of finding an S_y state (if you choose to measure the eigenstates of S_y, but as far as the up and down states are concerned their probabilities are constant: $\alpha^2\%$ chance to be up and $\beta^2\%$ chance to be down.

This is called "spin precession" and is every bit analogous to a spinning top. This type of analysis is a major part of the technology of MRI (magnetic resonance imaging otherwise known as NMR, nuclear magnetic resonance).

Typically the Hilbert Space of an object like the electron is a "tensor product" of the spatial/mechanical variables and the internal (spin) vari-

ables. In the previous example we pretended that the electron was fixed in space and only asked questions about its internal states. In general we could consider spin AND other dynamical variables like position.

$$|\Psi\rangle \in \mathcal{H}_{\text{mech}} \otimes \mathcal{H}_{\text{spin}}$$

and $|\Psi\rangle$ is a tensor product of these variables as $|\Psi\rangle = |\psi\rangle|\chi\rangle$

For example, if we were interested in position we would write

$$\psi(x, y, z, t)\vec{\xi}$$

where $\vec{\xi}$ is a two component spinor
viz,

$$\Psi(x,t) = \psi(x,t)\vec{\xi} = \begin{bmatrix} \psi_\uparrow(x,t) \\ \psi_\downarrow(x,t) \end{bmatrix}$$

Often times we write simply $\psi_s(x,t)$ where s indexes the up or down component of spin (or, if it's a higher spin object this index describes which m_ℓ component is being referred to)

This is how total angular momentum works for a single object. We have orbital angular momentum via the spatial variables and spin angular momentum via the internal spinor variables.

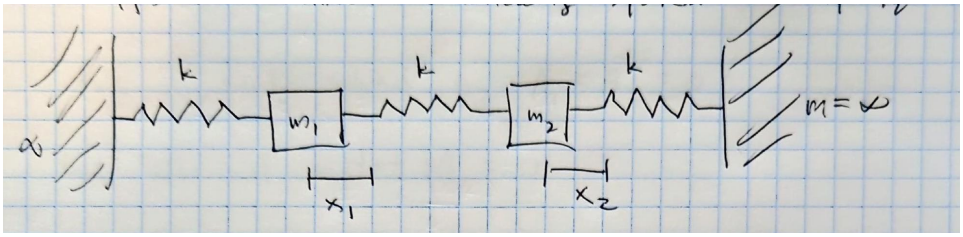
Normal Modes

Often times we encounter systems whose components are coupled thereby producing potentially MANY couple differential equations which must be solved. Here we will discuss a general method that works for a large class of coupled systems. Once we have this method we will be able to solve our linear chain model relatively easily as an eigenvalue, eigenvector equation. The solutions we obtain are called "Normal Modes" and a general solution is a linear combination of these normal mode solutions.

The normal modes behave as a basis for the vector space of solutions to our class of differential equations.

A Warm up example

Suppose we have the following system / $m_1 = m_2 = m$



we wish to solve for the motion $x_1(t), x_2(t)$

Newton says

we expect $x_i(t) = A_i e^{i\omega t}$ for both $x_1 \leqslant x_2$

plugging into Newton we have

$$-\omega^2 m A_1 = -2kA_1 + kA_2$$
$$-\omega^2 m A_2 = -2kA_2 + kA_1$$

this looks like a Matrix equation!

$$\begin{bmatrix} \omega^2 m - 2k & k \\ k & \omega_m^2 - 2k \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = 0$$

this is easily solved by $A_1 = A_2 = 0$! but also if $\det M = 0$

$$\left(\omega^2 m - 2k\right)^2 - k^2 = 0$$
$$\Rightarrow \omega_\alpha = \sqrt{\frac{k}{m}} \quad \text{or} \quad \omega_\beta = \sqrt{\frac{3k}{m}}$$

If we have frequency ω_α then $A_1 = A_2 = 1$. This means the two masses vibrate back and forth in the same direction.

If, however, we have frequency ω_β then $A_1 = -A_2$. This means the two masses vibrate back and forth in opposite directions. Try and visualize these two modes. The frequency of the β mode is $\sqrt{(3)}$ times faster than the α mode.

We call these two solutions "Normal Modes".

So we have found that in case α both masses oscillate in the same direction and in case β they move oppositely or 180° out of phase!

$$\vec{\eta}_\alpha = \begin{bmatrix} x_1^\alpha(t) \\ x_2^\alpha(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{i\omega_\alpha t}$$
$$\vec{\eta}_\beta = \begin{bmatrix} x_1^\beta(t) \\ x_2^\beta(t) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{i\omega_\beta t}$$

An arbitrary solution to our problem is a linear combination of these normal modes. The normal modes form a basis for the general solution to the motion of these coupled masses. Notice crucially (but not obviously) that we have no notion of translation symmetry of our system - this will be very important in our next example when we impose a periodicity on the system by arranging our masses and springs in a circle (or infinite chain).

$$\vec{x}(t) = \alpha\vec{\eta}_\alpha + \beta\vec{\eta}_\beta$$

where $\alpha, \beta \in \mathbb{R}$

General Solution To Normal Mode Problems

In a general problem in physics we often have a potential energy of the form

$$V = \sum_i \sum_j B_{ij} q_i q_j$$

We can write this in matrix language as

$$V = \vec{q}^T B \vec{q}$$

$$\begin{bmatrix} q_1 & q_2 & \cdots \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \cdots \\ B_{21} & \cdots & \\ \vdots & & \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \end{bmatrix} = \text{ some number}$$

where $\vec{q}$ represents the coordinates of each object in our system (i.e. $q_1, q_2, \cdots$) and B encodes all the various interactions between all the objects that make up our system. We shall generally say there are N objects in our system. This matrix is rarely diagonal. If it were diagonal it would generally be very easy to solve our problem. A diagonal B matrix would say that all the objects in the system don't interact with each other (there would only be "self-interactions" whatever that means). So typically we have some non-diagonal matrix.

Typically we consider only interactions between nearest neighbors so that often B is an off-diagonal or even tri-diagonal matrix.

For example, q_4 and q_5 interact and q_4 interacts with q_3 but q_3 and q_5 do NOT interact.

At any rate, the method we shall develop will work with ANY B no matter how complicated or unrealistic. The bulk of solving a system involves this potential energy matrix B.

Now, in an arbitrary system each object will also, in general, carry a kinetic energy.

$$T = \sum_i \sum_j A_{ij} \dot{q}_i \dot{q}_j$$

Where A can be completely general but is (almost?) always diagonal and contains the mass terms for each object in the system.

In matrix language this is

$$T = \vec{\dot{q}}^T A \vec{\dot{q}} = \dot{q}^T A \dot{q}$$

Where we will no longer write the little vector arrow for notational clarity. Just be mindful that all of the q's you will see are N dimensional vectors!

Therefore, we have a total energy

$$H = T + V = \frac{1}{2} \dot{q}^T A \dot{q} + \vec{q}^T B \vec{q}$$

Now, if energy is conserved then

$$\frac{d}{dt} H = 0$$

$$\Rightarrow 2\dot{q}^T (A\ddot{q} + Bq) = 0$$

which is only zero when

$$A\ddot{q} + Bq = 0$$

Now we shall restrict our attention to oscillatory systems and will seek solutions to this equation to have the form:

$$q = Xe^{i\omega t}$$

This is to say that every object of our system oscillates in time with frequency ω. Admittedly this doesn't seem very realistic at all. Sure there are situations where each object is bouncing back and forth at the same frequency but typically they might move at all sorts of different frequencies and phases! That is fine but for now we are only seeking the "normal modes" which oscillate at the same frequency. Once we find those the normal modes will form a basis of all possible vibrations!

So we plug $q = Xe^{i\omega t}$ into

$$A\ddot{q} + Bq = 0$$

and find

$$(B - \omega^2 A)X = 0$$

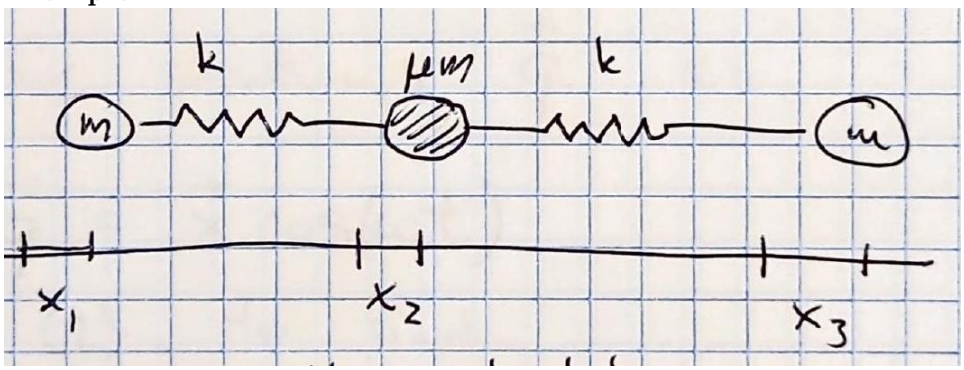
which is certainly zero if $X = 0$ or if

$$\det(B - \omega^2 A) = 0$$

We can solve this equation to find N frequencies ω_k and for each ω_k we will have a particular amplitude $X \to X_k$

What we have secretly done is we have diagonalized our system into a particular set of coordinates that simplifies our system into a set of N uncoupled equations.

Example



x_1, x_2, x_3 are the stretches from equilibrium

$$KE = T = \frac{1}{2}m\left(\dot{x}_1^2 + \mu\dot{x}_2^2 + \dot{x}_3^2\right)$$

$$PE = V = \frac{1}{2}k\left[(x_2 - x_1)^2 + (x_3 - x_2)^2\right]$$

$$A = \frac{1}{2}m\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad B = \frac{1}{2}k\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\dot{q}^T A\dot{q} = T$$

$$q^T Bq = V$$

the Hamiltonian/Energy is

$$\begin{aligned} H &= T + V \\ &= \dot{q}^T A\dot{q} + q^T Bq \end{aligned}$$

if energy is conserved then

$$\begin{aligned} &\frac{d}{dt}H = 0 \\ \Rightarrow &2\dot{q}^T(A\ddot{q} + Bq) \end{aligned}$$

$$A\ddot{q} + Bq = 0$$

we look for solutions q as

$$\vec{q} = \vec{x}\cos(\omega t)$$

plug into (1) to find

$$-\omega^2 A\vec{x} + B\vec{x} = \left(B - \omega^2 A\right)\vec{x} = 0$$

Then $\det\left(B - \omega^2 A\right) = 0$
Solving this yields our eigenvalues ω

$$B - \omega^2 A = \begin{bmatrix} 1-\omega^2 & -1 & 0 \\ -1 & 2-\mu\omega^2 & -1 \\ 0 & -1 & 1-\omega^2 \end{bmatrix}$$

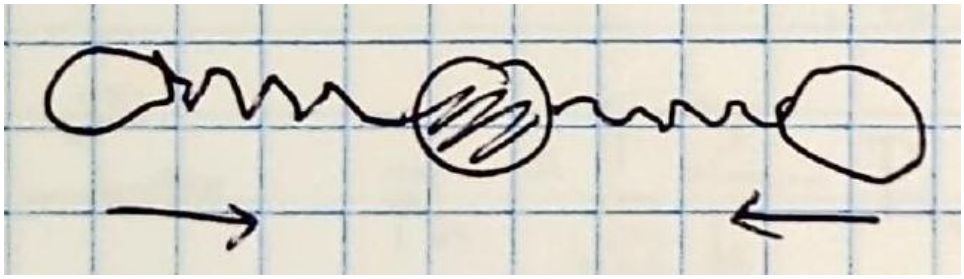
$$\Rightarrow \omega_1^2 = 0, \omega_2^2 = 1, \omega_3^2 = 1 + (2/\mu)$$

the eigenvectors of the eigenvalues are

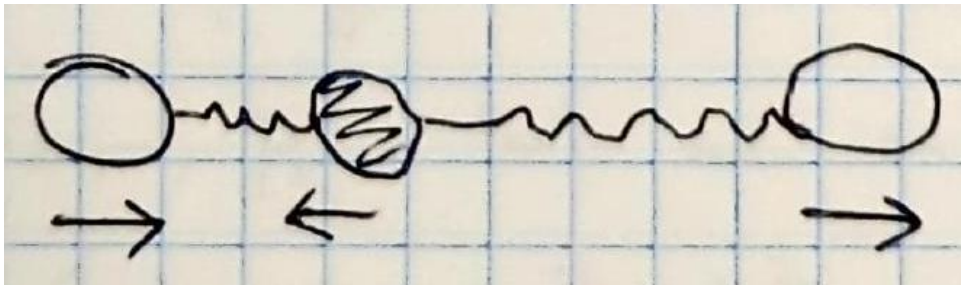
$$\vec{x}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \vec{x}_3 = \frac{1}{\sqrt{2+4/\mu^2}}\begin{bmatrix} 1 \\ -2/\mu \\ 1 \end{bmatrix}$$

Eigenvector $\vec{x}_1$ has zero frequency and each mass moves to the right. This is a simple translation: like the whole system is being gently carried around in space without vibrating anything.

Eigenvector $\vec{x}_2$ has the middle mass not moving and the outer masses vibrating in opposite directions.



Eigenvector $\vec{x}_3$ has outer masses moving together in the same direction but the inner mass moving opposite.



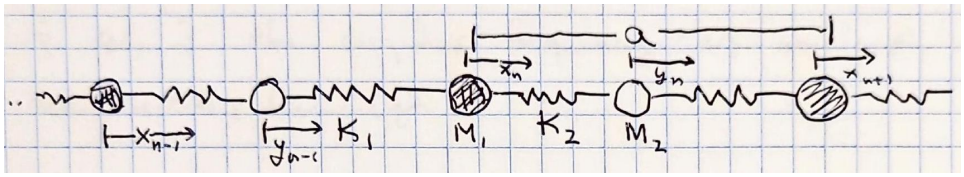
These are the thine Normal Modes, any complicated motion is a linear commination of these normal modes with their corresponding frequencies.

$$\vec{q}(t) = \sum_k \alpha_k \vec{x}_k \cos(\omega_k t)$$

Where α_k is a real number. This is the most general solution to the example above.

Classical Linear Chain Model

Now, we shall begin our quantum quest to discover quantized waves. We shall consider an entirely classical linear system of N coupled masses below. We shall imagine they are connected together in a VERY large circle such that we enforce periodic boundary conditions. Alternatively we could take N to ∞ but either way there will be a translational symmetry imposed.



We will have two types of mass M_1 and M_2 and two types of springs K_1 and K_2. You might think of this as a unit cell of some crystal of atoms (say, a single atomic layer of NaCl or a polymer).

Newton's 2^{nd} law says for each n we have

$$M_1\ddot{x}_n = K_2\left(y_n - x_n\right) - K_1\left(x_n - y_{n-1}\right)$$
$$M_2\ddot{y}_n = K_1\left(x_{n+1} - y_n\right) - K_2\left(y_n - x_n\right)$$

i.e. these equations govern the motion of M_1 and M_2 located at " n " viz x_n, y_n respectively. We could solve these coupled differential equations using methods from math classes or we can GUESS! Even little kids would reckon that the M_1 and M_2 will vibrate so let's guess general wave solutions for x_n and y_n

A general wave looks like

$$\begin{aligned}\xi(x,t) &= e^{ikx - i\omega t} \\ &= \cos(kx - \omega t) + i\sin(kx - \omega t)\end{aligned}$$

where k is the spatial frequency of the wave ω is the angular frequency. erg. wavelength and ovulation frequency
here our spatial variable where the ukases are located is an where $n \in \mathbb{N}$. In our case our position variables are an and the e^{ikan} originates from translational symmetry.

Before we solve this system in full generality let us consider a small and less complex version of this system and see how our machinery works in action:

Example
Consider a periodic chain of four equal masses connected with equal springs:

We have from Newton:

$$m\ddot{x_n} = K\left(x_{n-1} - x_n\right) + K\left(x_{n+1} - x_n\right)$$

We anticipate oscillatory solutions of this equation

$$x_n(t) = A_n e^{i\omega t}$$

assuming all masses oscillate in time as ω.

Writing ALL the equations in matrix form for each n we have (DO THIS!)

$$\begin{bmatrix} 2K - \omega^2 m & -K & 0 & -K \\ -K & 2K - \omega^2 m & -K & 0 \\ 0 & -K & 2K - \omega^2 m & -K \\ -K & 0 & -K & 2K - \omega^2 m \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = 0$$

$2K - \omega^2 m$ along the diagonal and $-K$ on the adjacent off-diagonals and $-K$ at the corners due to our self-imposed periodicity (the first and last masses are coupled).

Now, for periodic systems we have a symmetry; namely, translation by a.

If I apply a translation T to the circle of masses then whatever the motion I find should be invariant under this operation.

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

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

the eigenvectors of T are given by

$$T\vec{\xi} = e^{ika}\vec{\xi}$$

if

$$k = \frac{2j\pi}{aN}$$

where $j \in \mathbb{N}$ and aN is the length of the chain.

Notice the identity is given by T^N so that $e^{ikaN} = 1$.

Now, our equation can be written in terms of T as

$$(2K - \omega^2 m)\mathbb{1} - K(\hat{T} + \hat{T}^{-1})\vec{A} = 0$$

and we find,

$$x_n(t) = A_k e^{i(kan - \omega_k t)}$$

$$k = \frac{2j\pi}{aN}$$

where $j \in \mathbb{N}$

Alternatively we could just find the eigenvalues and eigenvectors of our whole matrix without our appeal to translational symmetry. The e^{ika} pops out when you find the eigenvalues/eigenvectors but it's easier to appreciate

it as a consequence of periodicity and will be VERY important when we discuss a periodic chain of quantum objects like electrons!

Ok, so use your favorite software (maple, matlab, mathematica, etc) to find the eigenvalues and eigenvectors of the matrix M

$$M = \begin{bmatrix} 2K - \omega^2 m & -K & 0 & -K \\ -K & 2K - \omega^2 m & -K & 0 \\ 0 & -K & 2K - \omega^2 m & -K \\ -K & 0 & -K & 2K - \omega^2 m \end{bmatrix}$$

The we solve $MA = 0$ and $\det M = 0$ to find

$$\omega_k = \sqrt{\frac{4K}{m}} \sin\left(\frac{ka}{2}\right)$$

With corresponding eigenvectors A_k (find these!).

Full generality

Back to the general system of N pairs of masses. We have to solve

$$M_1 \ddot{x}_n = K_2 (y_n - x_n) - K_1 (x_n - y_{n-1})$$
$$M_2 \ddot{y}_n = K_1 (x_{n+1} - y_n) - K_2 (y_n - x_n)$$

We pick candidate solutions as

$$x_n(t) = x_0 e^{i(kan - \omega t)}$$

$$y_n(t) = y_0 e^{i(kan - \omega t)}$$

Plugging these into our Newtons 2^{nd} Law we Find

$$-\omega^2 M_1 x_0 = K_2 (y_0 - x_0) - K_1 \left(x_0 - y_0 e^{-ika}\right)$$
$$-\omega^2 M_2 y_0 = K_1 \left(e^{ika} x_0 - y_0\right) - K_2 (y_0 - x_0)$$

Let's consider the case $k = 0$ (we could do it all at once but this route is illuminating). This is an infinite wavelength wave where all the masses are stretched by the same amount oscillating at the same frequency

if $k = 0$ then we can write this as a matrix equation

$$\begin{bmatrix} -\frac{K_1+K_2}{M_1} + \omega^2 & \frac{K_1+K_2}{M_1} \\ \frac{K_1+K_2}{M_2} & -\frac{K_1+K_2}{M_2} + \omega^2 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We find solutions from our determinant equation

$$-(K_1 + K_2)(M_1 + M_2)\omega^2 + M_1 M_2 \omega^4 = 0$$

from which we find

$$\omega_1 = 0$$

$$\omega_2 = \sqrt{\frac{(K_1 + K_2)(M_1 + M_2)}{M_1 M_2}}$$

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

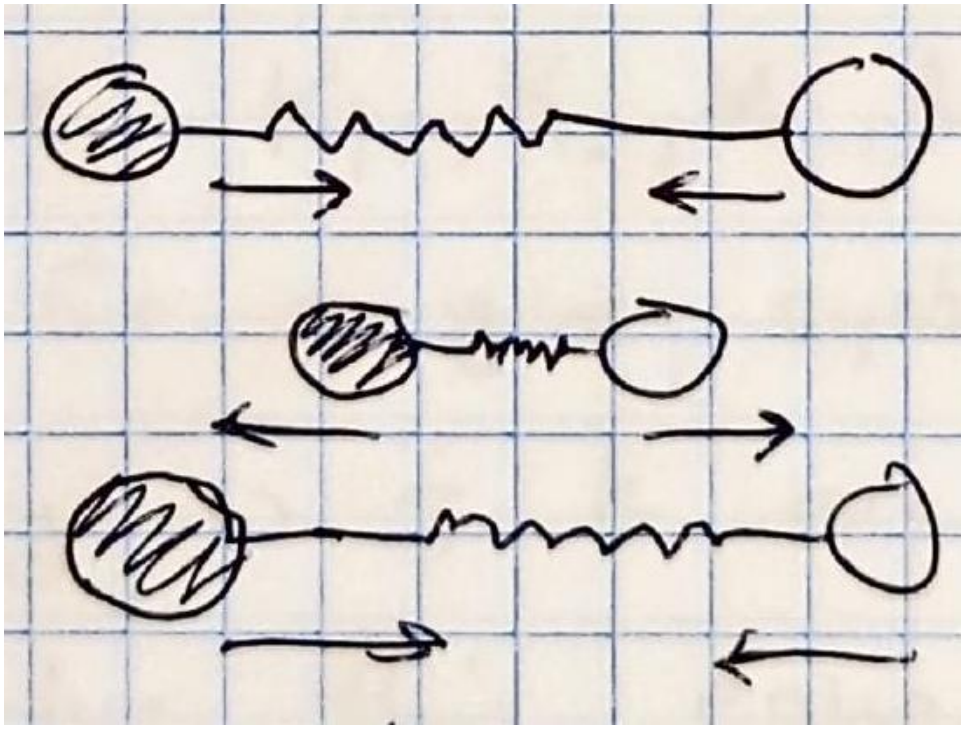
$$\vec{x}_2 = \begin{bmatrix} -M_2/M_1 \\ 1 \end{bmatrix}$$

What do these solutions mean?

The first has zero frequency and both masses translate to the right. Nothing is vibrating and the whole chain is simply translating in space! Boring! (these types of solutions typically show up - imagine holding a rope vertically and slowly walking with it: all the masses translate with no vibration).

The second solution is more interesting

the x's and y's move in opposite directions! e.g.



This is a kind of "beating" motion since each pair of masses are beating in the same way in opposite directions. There are no waves propagating through the whole system yet as we expect since $k = 0$ or $\lambda \to \infty$

What about $k \neq 0$?

Then we must solve

$$\begin{bmatrix} \frac{-K_1+K_2+\omega^2}{M_1} & \frac{K_1 e^{-ika}+K_2}{M_1} \\ \frac{K_1 e^{ika}+K_2}{M_2} & \frac{-K_1+K_2+\omega^2}{M_2} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Again we solve for eigenvalues and eigenvectors

$$-(K_1 + K_2)(M_1 + M_2)\omega^2 + M_1 M_2 \omega^4 + 2K_1 K_2(1 - \cos ka) = 0$$

the $\cos(ka)$ makes ALL the difference!

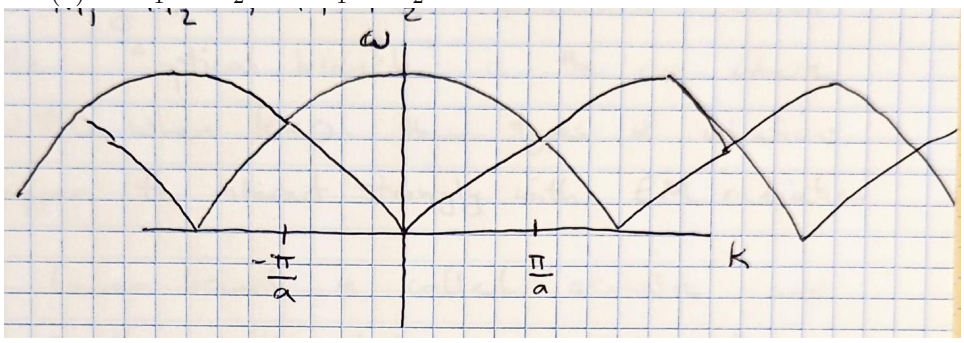
Take stock here. K_1, K_2, M_1, M_2 and a are all fixed values for a given system. They're not variables; only k and ω are variables so if we solve this equation we will get some function

$$\omega(k)$$

This is called a "dispersion relation". It connects spatial and temporal frequencies. It is very often the key object of study in materials science and condensed matter physics.

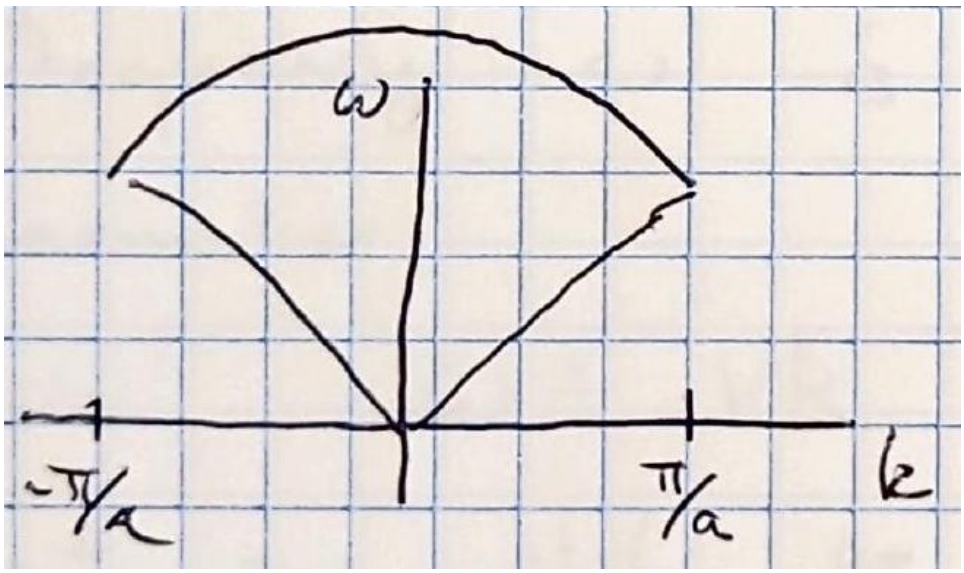
Let's plot some various cases

(1) $M_1 = M_2$ and $K_1 = K_2$



Usually we only draw the "zone" from $-\frac{\pi}{a} \to \frac{\pi}{a}$ since it is periodic. This is called the Brillion zone

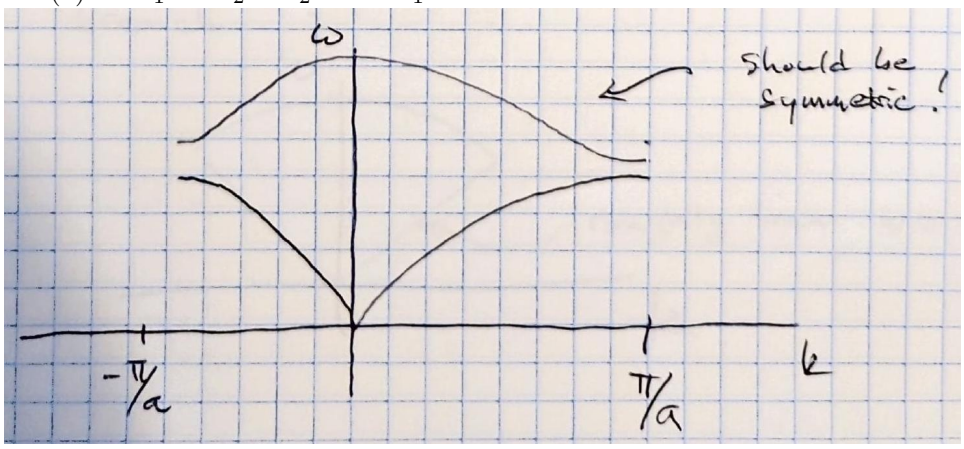
viz



We say there are two "branches" of this system. The one where $\omega \neq 0$ when $k = 0$ is called the "optical branch". These types of vibrations happen to interact strongly with EM radiation and represent the beating/oscillation between the two types of mass. The lower branch is called "acoustic" and is similar with what we typically associate with sound.

For example, a garden variety sound wave has frequency ω proportional to the wave number k as $\omega = vk$ where v is the speed of sound and is given by the slope of the acoustic branch near $k = 0$.

(2) $M_1 = M_2 \quad K_2 = 1.5K_1$



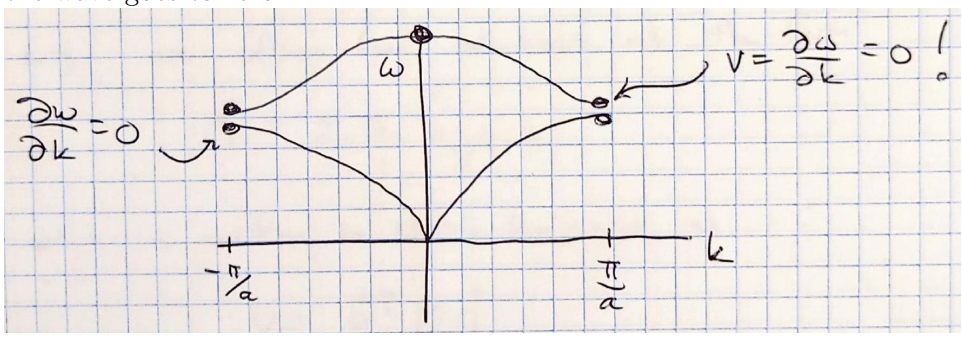
Notice that a gap opens up when $K_1 \neq K_2$. This is an important effect in materials science that we won't get into here only to say that the speed of these waves is identically zero at the zone boundaries. They are standing waves when $k = \frac{\pi}{a}$

The "group velocity" of a general wave is given by

$$\frac{\partial w}{\partial k} = v$$

if $\omega = vk$ then $\frac{\partial \omega}{\partial k} = v$ as above.

However, the group velocity is not required to vanish as ω or $k \to 0$, it is only the places where the slope of the dispersion is zero that the speed of the wave goes to zero.



For our general case here we have

$$\frac{\partial\omega}{\partial k} = \frac{2K_1K_2a\sin(Ka)}{2\left(K_1+K_2\right)\left(M_1+M_2\right)\omega - 4M_1M_2\omega^3}$$

When $k = \pm\pi/a$ then $\sin\left(\frac{\pi}{a}a\right) = 0$
and indeed standing waves appear at $\pm\pi/a$ as we expect.

Now, this analysis has been for a single value of n. However, by translational symmetry all other values of n get scaled by e^{ikan}

$$x_n(t) = x_{0k}e^{i(kan-\omega(k)t)}$$

$$y_n(t) = y_{0k}e^{i(kan-\omega(k)t)}$$

These represent solutions for each of the masses. The total general solution is a linear combination of the sum of all of these wave solutions. Usually, we're only interested in the normal modes because they form a basis but also because they're easy to visualize. It's best to view some animations of these normal modes on YouTube or, better yet, build a computer program to animate them yourself!

Now we shall apply quantum mechanics to this system. We shall take a slightly different path than above and it will be VERY difficult and non-obvious. Abandon all hope ye who enter here. When we make it to the other side of Styx we shall find a beautiful and deep simplification. Let's go.

Particles and Waves

Recall that $\psi(\vec{x})$ assigns a $\mathbb{C}$-number to all points in a space such that

$$\int_{-\infty}^{\infty}|\psi|^2dx = 1 \text{ or } \int\psi^*(x)\psi(x)dx < \infty$$

Functions that behave this way can generally be composed of plane waves as follows

$$\psi(x) = \int\frac{dk}{\sqrt{2\pi}}\tilde{\psi}(k)e^{ikx} \to \frac{1}{\sqrt{2\pi}}\sum_k\tilde{\psi}_ke^{ikx}$$

where

$$\widetilde{\psi}(k) = \int\frac{dx}{\sqrt{2\pi}}\psi(x)e^{-ikx} \to \frac{1}{\sqrt{2\pi}}\sum_x\psi_xe^{-ikx}$$

where $\tilde{\psi}(k)^2$ represents the probability of an object having wave-number k (or $p = \hbar k$)

we previously studied the single harmonic oscillator using

$$\tilde{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \quad and \quad [\hat{x}, \hat{p}] = i\hbar\mathbb{1}$$

this was solved using

$$\hat{a} = a = \tfrac{1}{\sqrt{2}}(\hat{x} + i\hat{p})$$
$$\hat{a}^\dagger = a^\dagger = \tfrac{1}{\sqrt{2}}(\hat{x} - i\hat{p})$$

such that $\hat{H}$ becomes

$$\hat{H} = \left(a^\dagger a + \frac{1}{2}\right)\hbar\omega = \left(\hat{n} + \frac{1}{2}\right)\hbar\omega$$

then

$$\hat{H}|n\rangle = E_n|n\rangle$$

where $|n\rangle$ are the eigenstates of $\hat{H}$ with eigenvalue given by

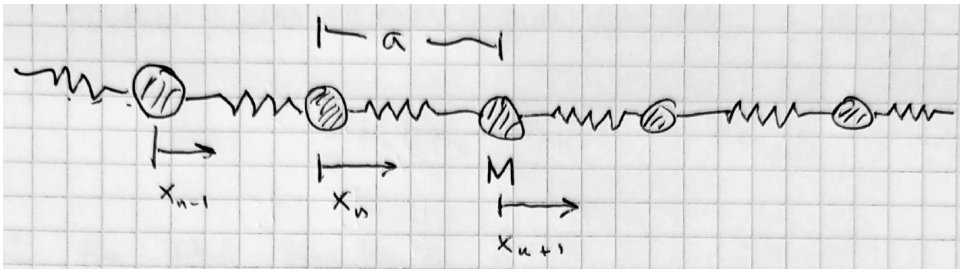
$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega$$

the ground state has $n = 0$ with

$$E_0 = \frac{\hbar\omega}{2} \text{ with state } |0\rangle$$

$|0\rangle$ is often called the "vacuum state". The vacuum has a non-zero energy implying that a harmonic oscillator can never be "at rest" in the way we usually think of classical systems. Classically a pendulum can be at rest in its lowest energy state.

Now, let's consider the following:



We have MANY identical quantum objects connected by identical springs of strength K. What are the quantum states of this complex system?

First, note that these don't have to be physical springs because ANY binding potential can be approximated by small deviations near the ground state. That is, via Taylor's expansion

$$V(x) = V(x_0) + V'(x)\big|_{x_0} dx + \frac{1}{2}V''(x)\bigg|_{x_0} dx^2 + \dots$$

At the minimum of a potential $V'(x)|_0 = 0$

$$\therefore \quad V(x) \approx \text{ constant } + \frac{1}{2}K(x - x_0)^2 + \theta\left(\Delta x^3\right)$$

for small Δx then $\Delta x^3 \approx 0$

What is the Hamiltonian? Here we will have many identical masses which each contribute a kinetic energy

Furthermore, each pair contributes a spring potential energy

$$\frac{1}{2}K(x_n - x_{n-1})^2 = PE_{\text{pair}}$$

$\therefore$ the total Hamiltonian is

$$H = \sum_n \frac{p_n^2}{2M} + \frac{1}{2}K(x_n - x_{n-1})^2$$

How the hell can we study this thing?! Classically we expect "waves" to propagate. We might try the following ansatz:

$$x_n(t) = x_0 e^{i(\text{kan} - \omega_k t)}$$

where, upon plugging in this ansatz, we find $\omega_k = \omega_0 \sin(ka/2)$ where k is the wave number. Then the most general solution is, a la Fourier, a linear combination of each "normal mode". How do we get there?

Details of our chain sum

$$H = \sum_n^N \frac{p_n^2}{2m} + \frac{1}{2}K(x_n - x_{n-1})^2$$

Goal: We aim to rewrite this expression in such a way that would allow us to solve it "simply".

The plan is to rewrite this Hamiltonian in a way that when we Fourier transform our variables many cancellations will occur.

$$x_n \xrightarrow{\mathscr{F}} x_n = \frac{1}{\sqrt{N}} \sum_k x_k e^{ikan}$$
$$p_n \xrightarrow{\mathscr{F}} \quad p_n = \frac{1}{\sqrt{N}} \sum_k p_k e^{ikan}$$

These are the discrete Fourier transforms of x_n, p_n where "an" is our discrete position variable. We can plug in these expansions for x_n, x_{n-1}, p_n. We shan't plug these in immediately: it would be a terrible mess. By the grace of folks MUCH more clever than myself, first, some tricks:

- Trick 1

$$(x_n - x_{n-1})^2 = x_n^2 - 2x_n x_{n-1} + x_{n-1}^2$$

Because we are summing over all n then $\sum_n x_n^2 = \sum_n x_{n-1}^2$
So inside the sum we can just combine these terms,

$$\sum (x_n - x_{n-1})^2 = \sum_n \left(2x_n^2 - 2x_n x_{n-1}\right)$$

- Trick 2
  write $2x_n x_{n-1} = x_n x_{n-1} + x_n x_{n-1}$

  Now we notice that x_n and x_{n-1} are neighbors as are x_n and x_{n+1}
  Since we're summing over ALL n we are free to write (inside the sum)

$$2x_n x_{n-1} = x_n x_{n-1} + x_n x_{n+1}$$

  Keep an eye on those indices!
  Now we have,

$$H = \sum_n \frac{p_n^2}{2M} + \frac{1}{2} K \left(2x_n^2 - x_n x_{n-1} - x_n x_{n+1}\right)$$
$$= \sum_n \frac{p_n^2}{2M} + K \left(x_n^2 - \frac{1}{2}\left(x_n x_{n-1} + x_n x_{n+1}\right)\right).$$

  Now the tricks become a bit harder but I hope the logic will be clear

- Trick 3
  write

$$\sum x_n^2 = \sum_{n'} \sum_{n} x_n x_{n'} \delta_{nn'}$$

$$\delta_{nn'} = \begin{cases} 1 \text{ when } n = n' \\ 0 \text{ when } n \neq n' \end{cases}$$

Try to convince yourself by writing out the terms of the sum for a few n, n'. This is a relatively common trick in vector calculus and tensor calculus - it's like multiplying by a matrix identity.

We also use this same trick for the other terms,

$$\sum_n x_n x_{n\pm 1} = \sum_n \sum_{n'} x_n x_{n'} \delta_{n\pm 1, n'}$$

$$\delta_{n\pm 1, n'} = 1 \text{ when } n \pm 1 = n'$$

I know, it's a BIG unholy mess but now our original H now looks like

$$H = \sum_n \frac{p_n^2}{2M} + K \sum_{n'} \sum_n x_n x_{n'} \delta_{nn'} - \frac{1}{2} \left(x_n x_{n'} \delta_{n-1,n'} + x_n x_{n'} \delta_{n+1,n'} \right)$$

We can "factor out" $x_n x_{n'}$

$$H = \sum_n \frac{p_n^2}{2M} + K \sum_{n'} \sum_n x_n x_{n'} \left[\delta_{nn'} - \frac{1}{2} \delta_{n-1,n'} - \frac{1}{2} \delta_{n+1,n'} \right]$$

- Trick 4
  The Fourier transform of $\delta_{nn'}$ is given by

$$\delta_{nn'} = \frac{1}{N} \sum_k e^{ika(n-n')}$$

and we also have, from Fourier,

$$x_n = \frac{1}{\sqrt{N}} \sum_k x_k e^{ikan}$$

$$p_n = \frac{1}{\sqrt{N}} \sum_k p_k e^{ikan}$$

We plug ALL this into H. The momentum part is

$$H_p = \sum_n \frac{p_n^2}{2M} = \frac{1}{2M} \sum_n \frac{1}{\sqrt{N}} \left(\sum_k p_k e^{ikan} \right) \left(\frac{1}{\sqrt{N}} \sum_{k'} p_{k'} e^{ik'an} \right)$$

$$H_p = \frac{1}{2MN} \sum_n \sum_k \sum_{k'} p_k p_{k'} e^{ian(k-k')}$$

But do you remember our Fourier transform of $\delta_{nn'}$?

$$\text{if} \quad \delta_{nn'} = \frac{1}{N} \sum_k e^{ika(n-n')}$$

then the same formula also applies if we change the letters of our indices (when summing over an index you can call it anything you want!). Say,

$$\delta_{kk'} = \frac{1}{N} \sum_n e^{ina(k-k')}$$

$$H_p = \frac{1}{2M} \sum_k \sum_{k'} p_k p_{k'} \cdot \sum_n \frac{1}{N} e^{ina(k-k')}$$

$$\therefore H_p = \frac{1}{2m} \sum_k \sum_{k'} p_k p_{k'} \delta_{kk'}$$

So that

$$H_p = \frac{1}{2M} \sum_k p_k^2$$

Wow! We find the momentum term has exactly the same form whether in terms of spatial variables or wavevector variables. This typically does not happen! Kronecker Deltas and Fourier sure are awesome! All that work for such a simple expression that doesnt look any different except for the index.

Maybe, just maybe, we will get something nice out of the potential energy term as well!

$$V = K \sum_n \sum_{n'} x_n x_{n'} \left[\delta_{nn'} - \frac{1}{2} \delta_{n-1,n'} - \frac{1}{2} \delta_{n+1,n'} \right]$$

Only one way to find out: plug and chug

$$V = K \sum_n \sum_{n'} \left(\frac{1}{\sqrt{N}} \sum_k x_k e^{ikan} \right) \left(\frac{1}{\sqrt{N}} \sum_k e^{ikan'} x_k \right) \left[\delta_{nn'} - \frac{1}{2}\delta_{n-1,n'} - \frac{1}{2}\delta_{n+1,n''} \right]$$
$$= \frac{K}{N} \sum_a \sum_{a'} \sum_k x_k^2 e^{ika(n+n')} \left[\delta_{nn'} - \frac{1}{2} \left[\delta_{n-1,n'} + \delta_{n+1,n'} \right] \right]$$

Now all the δ' s

$$(\cdots)[\frac{1}{N} \sum_k e^{ika(n-n')} - \frac{1}{2N} \sum_k e^{ika(n-1-n')} + \frac{1}{2N} \sum_k e^{ika(n+1-n')}$$

yikes!
Notice, however, that we can factor out $e^{ika(n-n')}$
viz,

$$(\cdots) \left[\frac{1}{N} \sum_k e^{ika(n-n')} \left(1 - \frac{1}{2} e^{-ika} - \frac{1}{2} e^{ika} \right) \right]$$
$$V = \frac{K}{N} \sum_n \sum_{n'} \sum_k x_k^2 e^{ika(n+n')} e^{ika(n-n')} \left[1 - \left(\frac{e^{ika} + e^{-ika}}{2} \right) \right]$$

Dear reader, someday you will see $\frac{e^{ikn} + e^{-ika}}{2}$ and exclaim triumphantly COSINE!!

Now, we can tidy up using

$$\sum_n e^{ika(n-n')} = \delta_{nn'}$$

Then,

$$V = K \sum_{n'} x_k^2 e^{ika(n+n')} \delta_{nn'} (1 - \cos ka)$$

Finally

$$V = K \sum_k x_k^2 (1 - \cos ka)$$

but

$$1 - \cos ka = 2\sin^2\left(\frac{ka}{2}\right)$$

$$= 2\left(\frac{\omega_k}{\omega_0}\right)^2$$

Which defines ω_k

All together

$$H = \sum_k \frac{p_k^2}{2M} + \frac{1}{2}M\omega_k^2 x_k^2$$

Dear Reader, this is a BUNCH of Harmonic oscillators. one for each value of k! Quantum mechanically all we have to do is put little hats on the p's and x's!

Now, finally!

$$\hat{H} = \sum_k \frac{\hat{p}_k^2}{2M} + \frac{1}{2}M\omega_k^2 \hat{x}_k^2$$

each value of k is an oscillator and we know how to solve the harmonic oscillator in quantum mechanics!

$$\hat{H} = \sum_k \left(a_k^\dagger a_k + \frac{1}{2}\right)\hbar\omega_k$$

WOW!!

$$a_k = \frac{1}{\sqrt{2}}\left(\hat{x}_k + ip_k\right)$$
$$a_k^\dagger = \frac{1}{\sqrt{2}}\left(\tilde{x}_k - i\hat{p}_k\right)$$
$$\left[a_k, a_k^+\right] = \delta_{kk'}$$
$$a_k\left|N_k\right\rangle = \sqrt{N_k}\left|N_k - 1\right\rangle$$
$$a_k^\dagger\left|N_k\right\rangle = \sqrt{1+N_k}\mid N_k + 1\rangle$$

What are these $|N_k\rangle$ states?

They're analogous to a single oscillator whose energies are separated by

$$\frac{1}{2}\hbar\omega_k = \Delta E_k$$

They represent some number of energy quanta in a state.

Previously we thought of $|n\rangle$ as "level 1, level 2, level 3, etc"

But we can also take the view that these represent particles each of which has energy $\frac{1}{2}\hbar\omega_k$! For example, $|3\rangle$ is a state with "three particles" or "quanta" in it and has total energy $3\frac{1}{2}\hbar\omega_k$.

"k" is the "wave number" of the particle. We call these quantized particles "Phonons". They are the quanta of "sound" or vibration.

How'd we get there?

We wrote the total Hamiltonian ; leveraged a Fourier transform to "decouple" the total systems into a bunch of independent oscillators one for each value of k

This process is called "diagonalizing the Hamiltonian"

$$\hat{H} = \left(a_k^\dagger a_k + \frac{1}{2}\right)\hbar\omega_k$$

What is the space of states?

Each value of k has a whole spectrum of oscillator states. Let's write them out

$$k_1 \quad k_2 \quad k_3 \quad k_4 \quad \cdots$$

in k_1 we can have N_{k_1} quanta of total energy

$$E_{k_1} = \left(N_{k_1} + \frac{1}{2}\right)\hbar\omega_{k_1}$$

where

$$\omega_{k_1} = \omega_0 \sin\left(\frac{k_1 q}{2}\right)$$

the states for $k = k_1$ look like

$$|0\rangle_{k_1}, |1\rangle_{k_1}, \ldots.. = |N_{k_1}\rangle$$

but there is also a tower of states for $k_2; k_3$, and on and on for each value of k.

We write the total state as a product of each individual oscillator Hilbert space

$$\begin{aligned} &|N_{k_1}\rangle\,|N_{k_2}\rangle\,|N_{k_3}\rangle \ldots \left|N_{k_j}\right\rangle \\ &= \left|N_{k_1} N_{k_2} N_{k_3} \ldots N_{k_j}\right\rangle \end{aligned}$$

The vacuum state will have ALL $N_k = 0$

$$|Vac\rangle = |00000, \ldots, 0\rangle$$

with energy

$$E_0 = \sum_j \frac{\hbar\omega_0}{2}$$

Where j is the total number of distinct k's.

We can "create" any state we want from the vacuum using the creation operator $a_k^\dagger$

For example,

$$a_5^\dagger|\text{vac}\rangle = |000010\ldots.0\rangle$$

This is a $k = 5$ state in the entire space of states

or

$$a_{23}^\dagger a_6^\dagger a_5^\dagger|Vac\rangle =| \, 0000110\ldots10\ldots\rangle$$

and so on.

These states are call FOCK states and live inside Fock space and act as

$$a_{k_5}\,|N, N_2\cdots N_5\ldots N_j\ldots\rangle = \sqrt{N_5}\,|N_1N_2\cdots(N_5-1)\ldots\rangle$$

here an annihilation operator reduced the number of quanta in N_5 by one.

Each slot is an independent oscillator and each quantum is a little packet of energy equal to $\frac{\hbar\omega_k}{2} = E_k$

Inside a material there are tons of phonons all flying around. They pass right through each other in this model but if we added an an-harmonic term to the potential we would find that phonons can interact and "bounce" off each other and electrons could even be "paired" by phonons thereby creating conditions for superconductivity!

Now, in our model there are no electrons to be found. All we have are these big heavy masses M (atoms?) held together by a periodic potential and we find quantized oscillations which themselves behave as quantum objects!

Tunneling

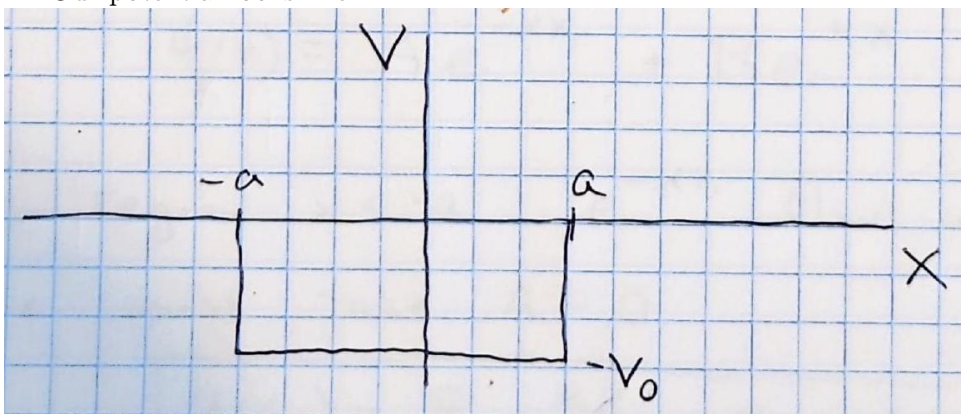
A finite potential well

A reasonable, yet simple, model for an electron in an atom is given by a finite potential energy well. If an election in an atom has enough energy then it will escape towards infinity and eventually hit a detector or be captured by some other potential well.

The full potential energy function of such a system is fairly complicated however we will extract the key elements by considering the following basic potential

$$V(x) = \begin{cases} -V_0 & -a < x < a \\ 0 & |x| > 0 \end{cases}$$

This is a finite version of the infinite well we had previously considered. Our potential looks like



It's essentially a hole in the ground. if you fall in, youve got to expend energy to climb back out. If we had a ball rattling around inside with total energy less than the potential depth then the ball will never escape. A ball must have energy greater than the depth of the well in order to escape. Pretty typical stuff. Let's see what quantum theory has to say.

Based on our previous work with the infinite potential well we anticipate sinusoidal-like solutions for energies close to $-V_0$. Interesting things, however, may happen due to the finite depth. For quantum objects moving towards the well from $\pm\infty$ perhaps they'll get trapped (a bound state) or perhaps they scatter of to infinity somehow (called a scattering state). Clearly this would all hinge on the total energy of the quantum object.

Let's study the different regions of our system under the assumption that $E < 0$ (i.e. the object is trapped)
Region I

For $x < -a$ the potential is zero. Schrödinger says,

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I}{dx^2} = E\psi_I$$

or

$$\frac{d^2\psi_I}{dx^2} = \kappa^2\psi_I$$

where $\kappa = \frac{\sqrt{-2mE}}{\hbar}$
since $E < 0$ then $\kappa \in \mathbb{R}$
Solutions to this differential equation are

$$\psi_I(x) = Ae^{-\kappa x} + Be^{\kappa x}$$

In our region $x < -a$ implies $e^{-\kappa x}$ blows up as $x \to -\infty$ so we must require $A = 0$

$$\psi_I(x) = Be^{\kappa x}$$

Region II

Now, in the region $-a < x < a$ we have $V(x) = -V_0$
So we must solve

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}}{dx^2} - V_0\psi_{II} = E\psi_{II}$$

or

$$\frac{d^2\psi_{II}}{dx^2} = -\ell^2\psi_{II}$$

where

$$\ell = \frac{\sqrt{2m\left(E + V_0\right)}}{\hbar}$$

here $E < 0$ but greater than $-V_0$ so that

$$l \in \mathbb{R}$$

Solutions to this type of equation are

$$\psi_{II}(x) = C\sin(lx) + D\cos(lx)$$

Region III

Finally for $x > a$ we have a similar situation to the first region and we require $\psi(x \to \infty) = 0$
so

$$\psi_{III}(x) = Fe^{-\kappa x}$$

Our problem is almost solved. We wound up solving three different differential equations; one for each specific region. However, we have yet to solve for the TOTAL wavefunction of this system.

To do so we must stitch these solutions together at the boundaries where each equation was solved. For example, $\psi_I(-a)$ should be equal to the solution of the second region $\psi_{II}(-a)$ as well as their derivatives at these boundaries.

We require

$$\psi_I(-a) = \psi_{II}(-a)$$

$$\psi_{III}(a) = \psi_{II}(a)$$

and

$$\left.\frac{d\psi_I}{dx}\right|_{-a} = \left.\frac{d\psi_{II}}{dx}\right|_{-a}$$

$$\left.\frac{d\psi_{II}}{dx}\right|_{a} = \left.\frac{d\psi_{III}}{dx}\right|_{a}$$

This amounts to mathematical surgery. We obtain the following equations:

$$Fe^{-\kappa a} = C\sin(la) + D\cos(la)$$

$$Be^{-\kappa a} = C\sin(-la) + D\cos(-la)$$

$$\kappa Be^{-\kappa a} = Cl\cos(-la) - Dl\sin(-la)$$

$$-\kappa Fe^{-\kappa a} = Cl\cos(la) - Dl\sin(la)$$

a

Four total equations and F, C, D, B are unknown, so that this is solvable. A more clever trick is to notice that since our potential is an even function then we can assume our solutions are either even or odd. i.e.

$$\psi(-x) = \pm\psi(x)$$

Let's focus only on the even solutions

$$\psi(x) = \begin{cases} Fe^{-xx} & x > a \\ D\cos(lx) & 0 < x < a \\ \psi(-x) & x < 0 \end{cases}$$

then at a

$$Fe^{-\kappa a} = D\cos(la)$$

and

$$-\kappa Fe^{-\kappa a} = -lD\cos(la)$$

Dividing these yields

$$\kappa = l\tan(la)$$

This is a formula for the possible energies of our system. viz

$$\frac{\sqrt{-2mE}}{\hbar} = \frac{\sqrt{2m\left(E+V_0\right)}}{\hbar}\tan\left[\frac{a\sqrt{2m(E+V_0)}}{\hbar}\right]$$

yikes!
Let's call $z = la$ and $z_0 = \frac{a}{\hbar}\sqrt{2mV_0}$ then notice that

$$\left(\kappa^2 + l^2\right) = \frac{2mV_0}{\hbar^2}$$

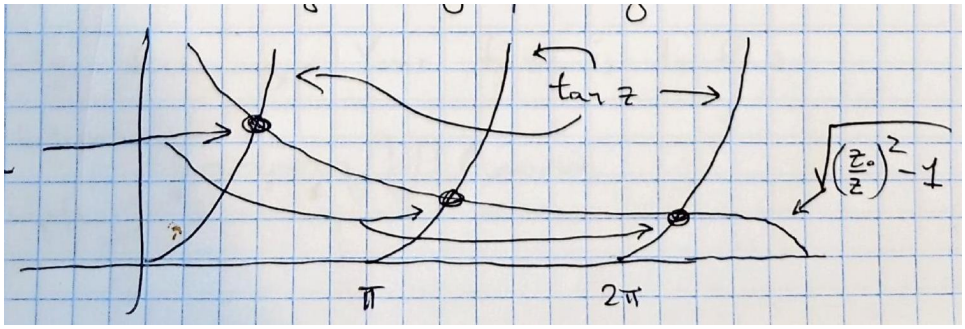
so that

$$\kappa a = \sqrt{z_0^2 - z^2}$$

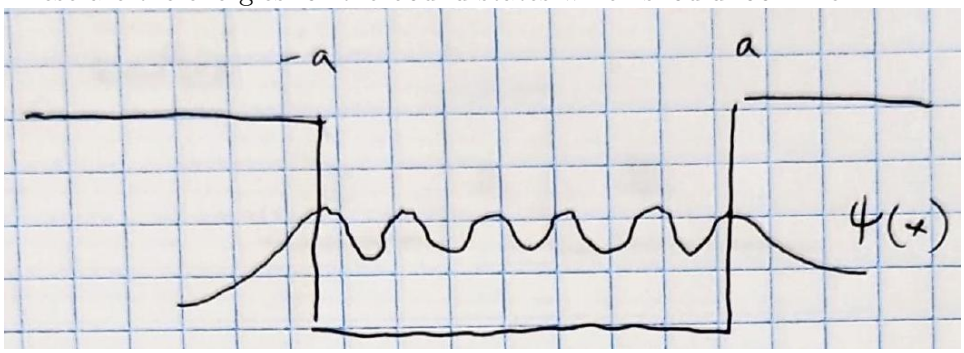
then

$$\tan z = \sqrt{\left(\frac{z_0}{z}\right)^2 - 1}$$

This is a very nasty transcendental equation which must be solved numerically or graphically. The possible energies are given below as the intersection points of these two functions.



These are the energies for the bound states which should look like

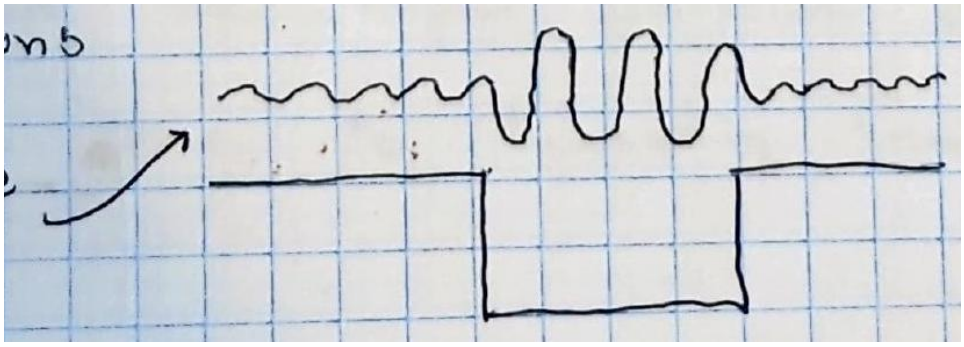


We notice that $\psi(x)$ "leaks" out of our well which means some times we will find it outside of the well! The probability that $x > a$ is non-zero! This is like putting a golf ball in a shoe box with no kinetic energy and sometimes seeing it outside of the shoebox! Surely this cant be true, right?!

This is called quantum tunneling and this effect HAS been observed in a variety of situations and experiments. In fact radioactivity depends on it!

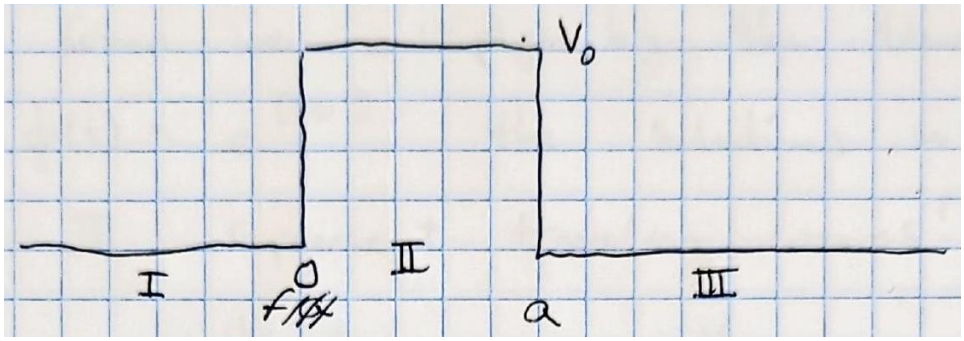
Now, what if we start with $E > 0$?
well, then $\kappa = \frac{\sqrt{-2mE}}{\hbar}$ is complex
and solutions will be wavelike rather than exponentially decaying outside of the well. Your task is to find these solutions

They would look like, for example,



A finite barrier

Now let's examine the situation whereby a quantum object impinges upon some potential barrier of height V_0 which is taken to be greater than the total energy of our quantum object. See below,



$$V = \begin{cases} 0 & x < 0 \\ V_0 & 0 < x < a \\ 0 & x > a \end{cases}$$

We shall proceed in exactly the same way: we solve a Schrodinger equation in each region of our system and then surgically stitch the solutions together at the boundaries.

Region I

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_I}{dx^2} = E\psi_I$$

$$\Rightarrow \quad \psi_I(x) = Ae^{ikx} + Be^{-ikx}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}$$

Region II

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}}{dx^2} + V_0\psi_{II} = E\psi_{II}$$

$$\Rightarrow \psi_{II}(x) = Ce^{-\kappa x} + De^{\kappa x}$$

where

$$\kappa = \frac{\sqrt{2m\left(V_0 - E\right)}}{\hbar}$$

Region III

This region is just like region I so that we have

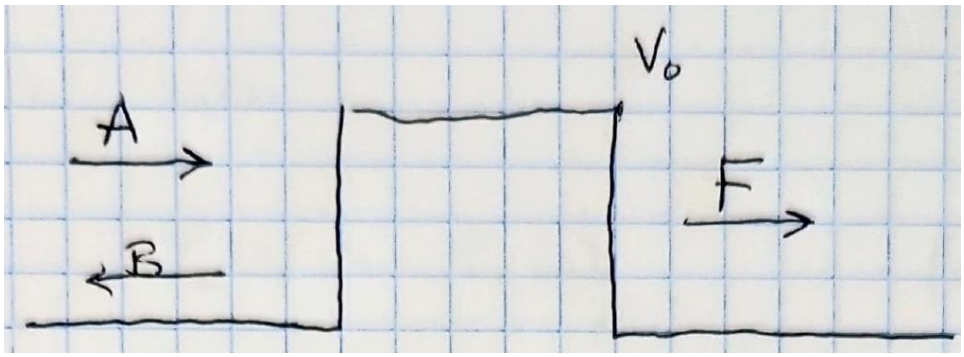
$$\psi_{\text{III}}\left(x\right) = Fe^{ikx} + Ge^{-ikx}$$

Notice that all we've done is solve the "time-independent" Schrödinger equation. The standard scenario is any time we want the time-dependent solution we multiply our spatial solution by $\phi(t) = e^{i\omega t}$. When we do this all of these standing waves become traveling waves! In region I the term $Ae^{ikx}e^{i\omega t}$ represents a traveling wave moving from $-\infty$ towards the barrier. a traveling wave moving to the right.

$Be^{-ikx}e^{i\omega t}$ represents a wave traveling towards the left from the barrier. A similar situation occurs in region III.

Let us solve a more narrowly focused problem where a quantum object comes in from the left, hits the barrier, and then reflects or transmits or does whatever we'll find it does.

A represents the incident wave, B represents the reflected wave, and F represents a transmitted wave (if that even can happen or not). We will set $G = 0$ since we dont anticipate any waves will be incident upon the barrier from $+\infty$



Now, we shall make some definitions:
Define:

$$R = \frac{|B|^2}{|A|^2}$$

is the relative probability the object will be reflected at $x = 0$. This is called the reflection coefficient. For example, if $B = 0$ there will be zero reflection and we also recall in quantum theory the square of the amplitudes represent probabilities.

Similarly, define

$$T = \frac{|F|^2}{|A|^2}$$

is called the transmission coefficient: the probability of transmission.
These are our two possible options so that

$$R + T = 1$$

to solve we must surgically stitch everything together using

$$\psi_{\mathrm{I}}(0) = \psi_{\mathrm{II}}(0) \tag{1}$$

$$\psi_{\mathrm{III}}(a) = \psi_{\mathrm{II}}(a) \tag{2}$$

$$\left.\frac{d\psi_I}{dx}\right|_0 = \left.\frac{d\psi_{\mathrm{III}}}{dx}\right|_0 \tag{3}$$

$$\left.\frac{d\psi_{\mathrm{III}}}{dx}\right|_a = \left.\frac{d\psi_{\mathrm{II}}}{dx}\right|_a \tag{4}$$

from (1)

$$A + B = C + D$$

from (2)

$$Fe^{ika} = Ce^{-\kappa a} + De^{\kappa a}$$

from (3)

$$Aik - Bik = -C\kappa + D\kappa$$

from (4)

$$Fike^{ika} = -\kappa Ce^{-\kappa a} + \kappa De^{\kappa a}$$

Four equations with five unknowns! We're sunk. However, we do have $T + R = 1$ and ultimately, for this analysis, were interested in T itself which is $T = \frac{|F|^2}{|A|^2}$

Now, recall the results from calculus that

$$\cosh x = \frac{e^x + e^{-x}}{2}$$
$$\sinh x = \frac{e^x - e^{-x}}{2}$$

also define $\gamma = \frac{\kappa}{k} - \frac{k}{\kappa}$

Then, plugging everything in and rearranging, we have

$$\frac{F}{A} = \frac{e^{-ika}}{\cosh(\kappa a) + i\frac{\gamma}{2}\sinh(\kappa a)}$$

Then

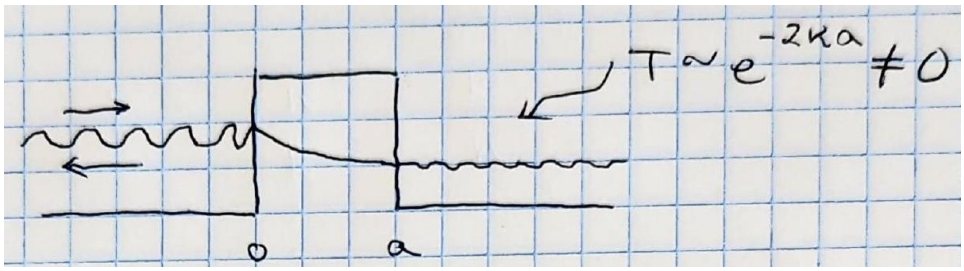
$$T = \frac{1}{\left(\cosh^2(\kappa a) + \frac{\gamma^2}{4}\sinh^2(\kappa a)\right)}$$

This is generally non-zero!

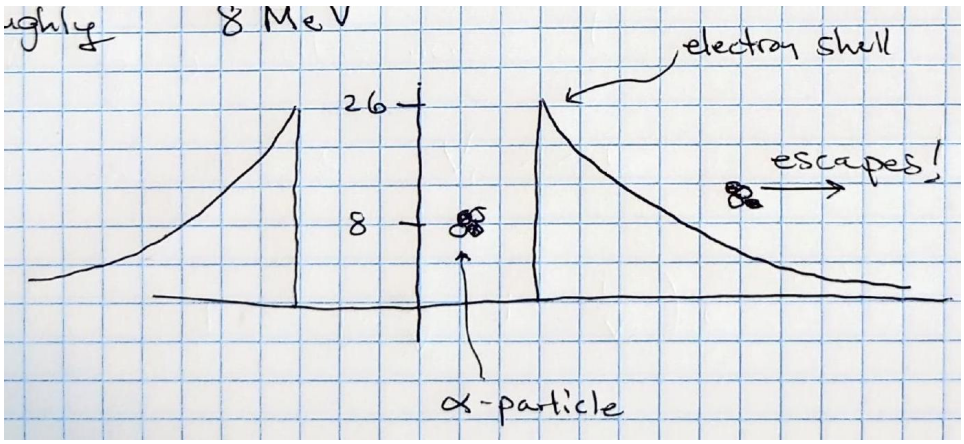
If we consider $V_0 \gg E$ and $a \gg 1$, i.e. a tall and wide barrier, then

$$T \approx 16\frac{E}{V_0}\left(1 - \frac{E}{V_0}\right)e^{-2\kappa a}$$

this $T \sim e^{-2ka}$ is one of the most famous results in quantum theory.



Gamow and Condon famously applied tunneling to the alpha radiation emitted by Polonium. Alpha particles are energetic helium nuclei (two protons and two neutrons). When a Polonium atom emits an alpha particle that alpha has to tunnel through Polonium's outer electron shells. Now, it was known that the α has energy around 8 MeV and the barrier it has to overcome is around 26 MeV. Gamow and Condon modeled the atom as a potential well with a Coulombic decay. See below



They considered that the α rattles around inside the well periodically being repelled by the Coulombic barrier. But based on our analysis every time the quantum α hits the barrier there is a small probability for transmission to occur. If this probability is large enough then sometimes you will find α particles emerging from the atom itself.

Now, this doesnt at all explain "why" Polonium atoms emit α particles (that requires new physics called the weak force). But what it does explain and predict is the half-life or decay constant of Polonium.

The tunneling model predicts a decay constant of around $\lambda = 0.3\mu s$ which is surprisingly close to the measured value of $\lambda = 0.26$ μs!

Hydrogen

A classical aside on central potentials

Recall, in our studies of , say, gravitation we found two primary symmetries and conservation laws. Noether's theorem gives a firm way of finding conservation laws if given a symmetry.

1. Conservation of Energy

$$\frac{dE}{dt} = 0$$

2. Conservation of Angular Momentum

$$\frac{d\vec{L}}{dt} = 0$$

for example,

$$\vec{F} = -\vec{\nabla} V(r)$$

and

$$m\ddot{r} = -\nabla V(r)$$

Let's consider

$$\vec{L} = \vec{r} \times \vec{p}$$

and note the following:

1. $\vec{r} \cdot \vec{L} = 0$

$$\vec{r} \cdot \vec{r} \times \vec{p} = 0$$

since $r \times p \perp r$

2. consider H in the x, y plane

$$H = KE + PE$$

$$H = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} m\dot{y}^2 + \frac{\alpha}{r}$$

$$x = r\cos\theta \quad \dot{x} = \dot{r}\cos\theta - r\sin\theta\dot{\theta}$$

$$y = r\sin\theta \quad \ddot{y} = \dot{r}\sin\theta + r\cos\theta\dot{\theta}$$

therefore

$$\dot{x}^2 = \dot{r}^2 \cos^2\theta + r^2\dot{\theta}^2 \sin^2\theta - 2r\dot{r}\dot{\theta}\sin\theta\cos\theta$$

and

$$\dot{y}^2 = \dot{r}^2 \sin^2\theta + r^2\dot{\theta}^2 \cos^2\theta + 2r\dot{r}\dot{\theta}\sin\theta\cos\theta$$

then, plugging in,

$$H = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{\alpha}{r}$$

Now

$$\frac{d}{dt}\vec{L} = \frac{d}{dt}(\vec{r}\times\vec{p}) = \dot{\vec{r}}\times\vec{p} + \vec{r}\times\dot{\vec{p}}$$

but

$$\dot{\vec{r}}\times\vec{p} = 0 \quad \therefore \quad \frac{d\vec{L}}{dt} = \vec{r}\times\dot{\vec{p}}$$

but $\dot{\vec{p}} = \vec{F}$

$$\therefore \quad \frac{d\vec{L}}{dt} = \vec{r}\times\vec{F}$$

but in a central potential $\vec{r}\|\vec{F}$

$$\therefore \frac{d\vec{L}}{dt} = 0.$$

Since $\vec{L}$ is conserved we will call its length ℓ. Then for a point mass we have

$$\ell = I\omega = I\dot{\theta} = mr^2\dot{\theta}$$

$$\Rightarrow \dot{\theta} = \frac{\ell}{mr^2}$$

then

$$H = \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r) = E$$

for some E. Then,

$$\frac{dH}{dt} = \frac{dE}{dt}$$

Back to Cartesian coordinates for simplicity and transparency. We wish to prove that $\dot{E} = 0$ which yields our conservation of energy law for this system.

$$\frac{dH}{dt} = \frac{d}{dt}\left(\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 + V(x,y)\right)$$

$$m\dot{x}\ddot{x} + m\dot{y}\ddot{y} + \frac{\partial V}{\partial x}\dot{x} + \frac{\partial V}{\partial y}\dot{y}$$

we have used the standard definition of the total derivative in multiple dimensions (viz, the chain rule).

$$\frac{d}{dt}V(x,y) = \frac{\partial V}{\partial x}\frac{dx}{dt} + \frac{\partial V}{\partial y}\frac{dy}{dt}$$

Note that V is a function of both x and y since r is given by Pythagoras accordingly.

but notice that

$$\frac{\partial V}{\partial x} = -F_x \qquad \frac{\partial V}{\partial y} = -F_y$$

$$\therefore \quad m\ddot{x}\dot{x} + m\ddot{y}\dot{y} - F_x\dot{x} - F_y\dot{y}$$

but $m\ddot{x} = F_x$ and $m\ddot{y} = F_y$

$$\therefore \quad \frac{dE}{dt} = 0$$

Notice crucially this proof had nothing to do with the form of V. All we needed was a conservative force.

Now, if

$$m\ddot{\vec{r}} = \frac{\alpha}{r^2}\hat{r}$$

we will study

$$\vec{L} \times m\ddot{\vec{r}}$$

It's not obvious why we should do this but we're able to sate our boredom by trying different things. Perhaps you wake up one morning and decide to

cross $\vec{L}$ with random things just to see what happens? As it turns out doing this manipulation does have more reason than that behind it, but this offers a nice example of how physics research can sometimes proceed.

To study this we need the following vector identity (prove this)

$$\vec{a} \times \vec{b} \times \vec{c} = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{b} \cdot \vec{c})\vec{a}$$

using this identity we have

$$\vec{L} \times m\ddot{\vec{r}} = -m\frac{\alpha}{r^2} r^2 \frac{d}{dt}\hat{r} = -m\alpha\dot{\hat{r}}$$

Well this looks funny but notice that $\frac{d\hat{r}}{dt}$ is only charging direction and not length here. This is simply $\dot{\vec{\theta}}$

So,

$$\vec{L} \times m\ddot{\vec{r}} = -\alpha m\frac{d\hat{r}}{dt}$$

But we can integrate this!

$$\int \vec{L} \times m\ddot{\vec{r}}dt = -\int \alpha m\frac{d\hat{r}}{dt}dt$$

but $\vec{L}$ is constant!

$$\vec{L} \times m\dot{\vec{r}} = -\alpha m\vec{r} - \vec{a}$$

where $\vec{a}$ is an integration constant (vector!). However,

$$\vec{L} \times m\dot{r}$$

$$\therefore \vec{p} \times \vec{L}$$

$$\therefore \vec{a} = \vec{p} \times \vec{L} - \alpha m\vec{r}$$

$\vec{a}$ is called the Runge-Lenz vector and it is conserved. Notice that the only reason this vector is conserved is because of the form of the potential $\frac{1}{r}$. This is entirely "accidental" or "dynamical" and therefore we say this is an "accidental/dynamical symmetry".

So here we get, for the Kepler problem, a symmetry and a conservation law and anytime we find conservation laws we should rejoice because they will

often allow us to solve complicated problems algebraically without leveraging complicated integrals and derivatives.

So we have,

$$\frac{d\vec{a}}{dt} = 0$$

It has to to be! It's an integration constant after all!

Maybe, just maybe, by the grace of nature this symmetry will hold in quantum land and we can use it to solve the hydrogen atom.

However, we might already see a problem: operators generally don't commute and we wrote

$$L \times p = -p \times L$$

Which form should we use for our quantum problems? This is a general difficulty with canonical quantization: anytime we have to choose an ordering of operators we dont a priori know which way to write them. What we do is write instead

$$AB = \frac{1}{2}(AB + BA)$$

For some bizarre reason this (almost) always works!

So if we put this ambiguity in by hand maybe the Quantum Runge-Lenz vector (called the Pauli-Lubansk) operator will be

$$\hat{\mathscr{L}} = \frac{1}{2m}(\hat{p} \times \hat{L} - \hat{L} \times \hat{p}) - \frac{\alpha}{r}\hat{r}$$

In this expression hats indicate operators. In the classical version we had a unit vector $\hat{r} = \frac{\vec{r}}{|r|}$ in the Lenz vector. So we were careful to replace it accordingly,

Therefore, we hope to solve

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

where

$$\hat{H} = \frac{\hat{p}^2}{2m} - \frac{\alpha}{\hat{r}}$$

again, $\hat{r}$ is the complete position vector and not the unit vector.

Now, quantum mechanically we expect angular momentum states due to rotational invariance of the coulomb problem. That is,

$$\hat{L}_z|l,m\rangle = m_\ell|\ell,m\rangle$$
$$L^2|\ell,m\rangle = \ell(l+1)|\ell,m\rangle$$

this is what we've learned about $SO(3)$

Also all our symmetries commute with $\hat{H}$. That's what it means to be a symmetry of a system. Therefore,

$$\left[\hat{H}, L_z\right] = 0$$

$$\left[\hat{H}, L^2\right] = 0$$

and given the hard work we've done we have a new accidental symmetry, namely

$$\left[\hat{H}, \hat{\mathscr{L}}\right] = 0$$

As usual we now wish to find the eigenstates and eigenvalues of $\hat{H}$,$\hat{L_z}$, and $\hat{\mathscr{L}}$. Since these operators all commute they share an eigen-basis we call $|\xi\rangle$. that is,

$$\hat{\mathscr{L}}|\xi\rangle = \mathscr{L}|\xi\rangle$$

and then these eigenvalues $\mathscr{L}$ will be part of our description of $\hat{H}$. As we will discover, they turn out to be related to the the principal quantum number n from chemistry class.

$$\hat{H}|n,l,m\rangle = E_n|n,l,m\rangle$$

Proof: Orbits Close

let $V(r)$ be a central potential e.g: $\frac{\alpha}{r}$. Newton says that for conservative forces we have

$$\sum \vec{F_i} = m\vec{a} \Rightarrow m\ddot{r} - mr\dot{\theta}^2 = -\frac{d}{dr}V(r)$$

and $L = mr^2\dot{\theta}$ is conserved. i.e.

$$\frac{dL}{dt} = 0 \Rightarrow mr^2\dot{\theta} = \text{ constant}$$

then, $\dot{\theta} = \frac{L}{mr^2} = \frac{d\theta}{dt}$ which gives us a different way to write time derivatives. viz,

$$\frac{d}{dt} \to \frac{L}{mr^2}\frac{d}{d\theta}$$

Let us rewrite Newton using this rescaling of time

$$\frac{L}{r^2}\frac{d}{d\theta}\left(\frac{L}{mr^2}\frac{dr}{d\theta}\right) = -\frac{2L^2}{mr^5}\left(\frac{dr}{d\theta}\right)^2 + \frac{L^2}{mr^4}\frac{d^2r}{d\theta^2}$$

$$\frac{L}{r^2}\frac{d}{d\theta}\left(\frac{L}{mr^2}\frac{dr}{d\theta}\right) - \frac{L^2}{mr^3} = -\frac{dV}{dr}$$

Now let's change variables yet again to $u = \frac{1}{r} \Rightarrow du = -\frac{1}{r^2}dr$
and find

$$\frac{du}{d\theta} = -\frac{1}{r^2}\frac{dr}{d\theta}$$

and

$$\frac{d^2u}{d\theta^2} = \frac{2}{r^3}\left(\frac{dr}{d\theta}\right)^2 - \frac{1}{r}\frac{d^2r}{d\theta^2}$$

Which yields

$$\frac{d^2u}{d\theta^2} + u = -\frac{m}{L^2}\frac{dV\left(\frac{1}{u}\right)}{du}$$

Now we consider a specific form for $V(r)$

Take $V(r) = \frac{\alpha}{r} = \alpha u$
then we must solve

$$\frac{d^2u}{d\theta^2} + u + \frac{\alpha m}{L^2} = 0$$

to find $u(\theta)$

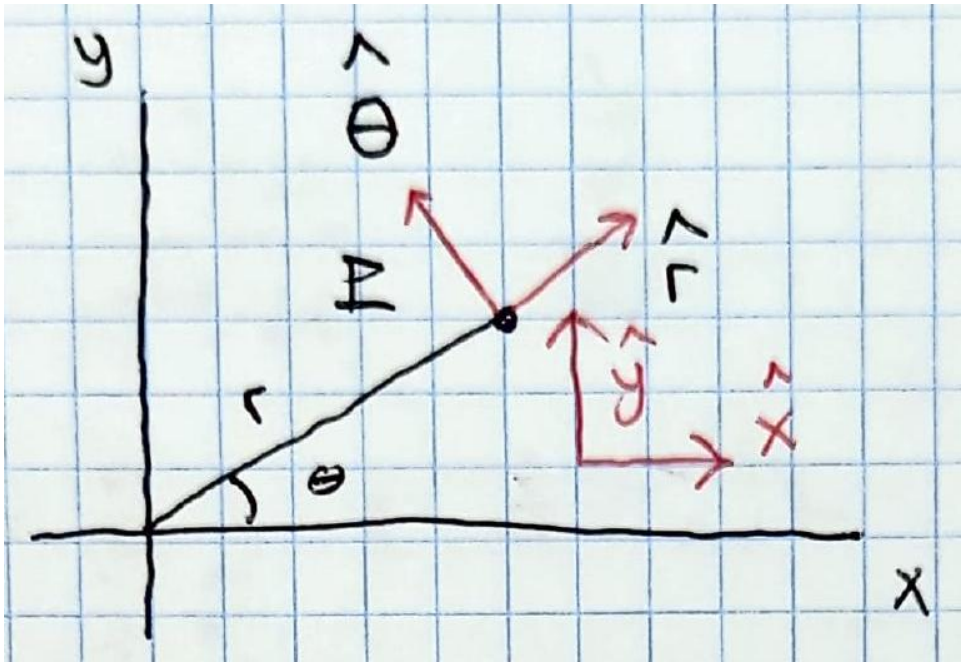
the solution to the homogenoues differential equation, $u'' + u = 0$, is

$$u(\theta) = \frac{1}{r} = -\frac{km}{L^2}\left[1 + e\cos\left(\theta - \theta_0\right)\right]$$

which is our standard solution for an elliptical orbit. notice that the orbit closes upon 2π rotation, viz, $u(0) = u(2\pi n)$ for $n \in \mathbb{N}$

Another way to Runge-Lenz and closed orbits
Consider the following:

1. $\dot{L} = 0 \Rightarrow$ motion in the plane



In terms of (x, y) we have

$$\hat{\theta} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad \hat{r} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

We will consider the spherical angle from the $\hat{z}$ axis to be $\phi = \pi/2$ (the equatorial plane).

Now

$$\ell = I\dot{\theta} = mr^2\dot{\theta}$$

and $\frac{1}{2}r^2\dot{\theta}$ is the area swept out in unit time (Kepler's Law). Let's call this $h/2$

$$\frac{r^2\dot{\theta}}{2} = \frac{\ell}{2m} = \frac{h}{2}$$

where the $1/2$ is just for convenience later.

2. Now we shall study motion in velocity space. E.g. we will use $\frac{d\vec{v}}{d\theta}$ to find an equation for $v(\theta)$. Note that

$$\frac{d\vec{v}}{d\theta} = \frac{d\vec{v}/dt}{d\vec{\theta}/dt} = \frac{\dot{\vec{v}}}{\dot{\vec{\theta}}}$$

but

$$\dot{\vec{v}} = \vec{a}$$

$$\dot{\vec{\theta}} = \frac{\ell}{mr^2}$$

but, we have,
$|\vec{a}| = -\frac{GM}{r^2}\hat{r}$ from Newton's gravity in the $\hat{r}$ direction
then

$$\frac{\dot{v}}{\dot{\theta}} = -\frac{GM}{r^2}\frac{mr^2}{\ell}\hat{r} = -\frac{GM}{h}\hat{r}$$

Now notice

$$\frac{d\hat{\theta}}{d\theta} == \frac{d}{d\theta}\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} -\cos\theta \\ -\sin\theta \end{bmatrix} = -\hat{r}$$

then

$$\frac{\dot{v}}{\dot{\theta}} = \frac{GM}{h}\frac{d\hat{\theta}}{d\theta}$$

Now pull out a $d/d\theta$ and we have

$$\frac{d}{d\theta}\frac{GM}{h}\hat{\theta} = \frac{d\vec{v}}{d\theta}$$

or

$$\frac{d}{d\theta}(\frac{GM}{h}\hat{\theta} - \vec{v}) = 0$$

which implies that $\frac{GM}{h}\hat{\theta} - \vec{v} = \vec{w}$ is a constant vector

$$\therefore \vec{v} = \vec{w} + \frac{GM}{h}\hat{\theta}$$

$\therefore$ we have an equation for $\vec{v}$ Which we know must lie in the x, y plane

What we want is $v(\theta)$. This would yield some curve in velocity space. But that's easy, we can use the dot product! By definition we have

$$v(\theta) = \vec{v} \cdot \hat{\theta} = r\dot{\theta}$$

but

$$r\dot{\theta} = \frac{h}{r}$$

and

$$\vec{v} \cdot \hat{\theta} = w \cos\theta + \frac{GM}{h} = v(\theta) = \frac{h}{r}$$

or

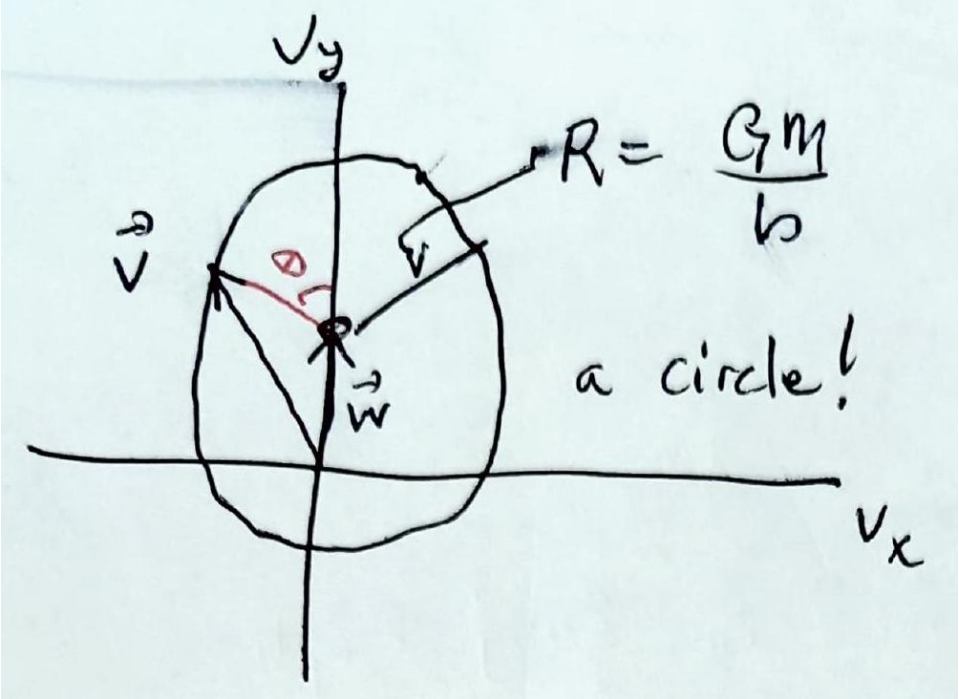
$$\frac{GM}{h}(1 + e\cos\theta) = \frac{h}{r}$$

where $w = \frac{GMe}{h}$. e is called the "eccentricity" of the orbit in position space.

But we get MORE! Back to $\vec{v} = \vec{w} + \frac{GM}{h}\hat{\theta}$ we have

$$\vec{v} - \vec{w} = \frac{Gm}{h}\hat{\theta} \Rightarrow (\vec{v} - \vec{w})^2 = \left(\frac{GM}{h}\right)^2$$

which implies we must have a circle in velocity space centered at $\vec{w}$ with radius $\frac{GM}{h}$!



Planets move in circles in velocity space (ALWAYS) but are conic sections in position space! The velocity space gives us circles with varying radii and position that translate directly to various conic sections in position space! This seems VERY deep.

Therefore it appears there may be some hidden symmetry behind $V(r) = \frac{k}{r}$ that not immediately clear.

It feels like $SO(3)$ in velocity space or something and we know we have $SO(3)$ in position space. Maybe (hypothesis!) there's some higher symmetry in phase space $(p_x, p_y, p_z; x, y, z) \in \mathbb{R}^6$ that behaves like

$$SO(3) \otimes SO(3)$$

or

$$SO(3) \oplus SO(3)$$

or perhaps even some higher symmetry group which has these symmetries as a subgroup. Maybe we can grope around in the dark and hit the right target? It seems practical to pick the simplest possibility (above) and pursue but there are TONS of groups which have these symmetries as a subgroup...we could be here all day everyday!

Okay, so now that we're convinced there may be some underlying symmetry at play how can we proceed?

Well, one option is to notice that we've done something similar by following Lie and his group theory.

The proper way to pursue this classically is to define a Poisson bracket on phase space and study the space of functions and yadda yadda yadda but we haven't developed those tools yet.

Instead we will pursue the idea in quantum land.

We wish to find the eigen-states of our symmetry but we don't know what the symmetry even is!

Previously, we said:

"if we have SO(3) $\Rightarrow$ we have angular momentum generators/states"

and we proceeded to solve for those generators by considering infinitesimal group transformation. once we had the matrices we invoked quantum theory to claim that "the eigenvectors of rotation generators are the states of definite angular momentum".

So, explicitly, we found states labeled by ℓ and m we called $|l, m\rangle$

such that

$$\hat{L}_z|\ell, m\rangle = m_\ell|\ell, m\rangle$$

and

$$\vec{L}^2|\ell, m\rangle = \ell(\ell+1)|\ell, m\rangle$$

where, for example,

$$\hat{L}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

or, for continuous functions (infinite dimensional vectors) in position basis,

$$\hat{L}_z = i\left(y\partial_x - x\partial_y\right)$$

and so forth.

In that method we used symmetry to find the operators.

But in our central potential we already have the operators!

to recap, we have:

1. conservation of angular momentum

$$\vec{L} = \vec{r} \times \vec{\rho}$$

2. conservation of the Runge-Lenz vector

$$\vec{a} = \vec{M} = \vec{p} \times \vec{L} - km\frac{\vec{r}}{|r|}$$

But we've seen $\vec{L}$ before. it obeys the $SO(3)$ (also $SU(2)$) Lie Algebra.

From here forward we shall assume we're dealing with operators and I will forgo explicitly writing the little hat on everything. If you get confused just put hats on stuff!

We have,

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k$$

and we said ANYTHING that obeys this algebra generates $SO(3)$ (or the $SU(2)$ double cover of $SO(3)$)

Will we also get $SO(3)$ when we ask

"what is $[M_i, M_j]$"

if so then we would find that M also generates an $SO(3)$ and we would have two separate $SO(3)$'s which would imply a total group of

$$SO(3) \oplus SO(3)$$

that is, an $SO(3)$ for M and L which don't interact. If we find M and L don't commute then we wouldn't have separate $SO(3)$'s and we would have more work to do to find how they interact with each other.

This sounds like we have a reasonable plan. It may be a lot of algebra but we can do it and maybe, just maybe, the group of symmetries of the central potential will jump out at us and make itself obvious. Wouldn't that be nice?

Before we proceed to checking these commutators we must prepare a couple observations

First, in order to be a symmetry of our system we had better have

$$[\vec{L}, \hat{H}] = 0$$

and

$$[\vec{M}, \hat{H}] = 0$$

Check this is the case (careful $AB \neq BA$) by using $[\vec{r}, \vec{p}] = i\hbar\mathbb{1}$ as always when doing quantum mechanics!

Secondly, in quantum land we never know whether to write

$$M \sim \vec{p} \times \vec{L} + km\vec{r}$$

or

$$M \sim -\vec{L} \times \vec{p} + kur\tilde{r}$$

Classically $\vec{p} \times \vec{L} = -\vec{L} \times \vec{p}$ but in quantum land order matters! We don't know which should be used and we will generally get different results using one or the other (try it)

So, the usual solution is that we will study the symmetrized form and hope for the best! I can hear you screaming in discontent at this proposition. At this point we have zero reason this will lead to realistic predictions but we can always desperately try something and if it doesn't work then it doesn't work and we try again!

So, we write

$$\vec{M} = \frac{1}{2m}(\vec{p} \times \vec{L} - \vec{L} \times \vec{p}) - \frac{k\vec{r}}{r}$$

since

$$\frac{a \times b}{2} + \frac{a \times b}{2} = \frac{a \times b - b \times a}{2}$$

Noticed I divided out m from the potential term.

Well, now we're ready for our plan. Let's find $[M_i, M_j]$

First, we notice that ultimately we will need commutators like

$$[L_i, r_j]$$

$$[L_i, p_j]$$

since $[r_i, p_j] = i\hbar\delta_{ij}$ and $\vec{M}$ contains these operators.

There's a really handy way to derive vector relations using the components of vectors and various standard matrices like Levi-Civita (ϵ_{ijk} and Kronecker (δ_{ij}). Cross products may be written as

$$\vec{a} = \vec{b} \times \vec{c} \Rightarrow a_k = (\vec{b} \times \vec{c})_k = a_i b_j \epsilon_{ijk}$$

Here we implicitly sum over the repeated indices ij using the totally anti-symmetric Levi-Civita tensor (density). Notice further that everything in this formula is a number: no vectors or cross products in sight. This expression gives us formulas for the components of vectors and are especially handy for deriving vector identities (though, it takes practice).

Now, we can study our commutators using these types of expressions. First, (keeping in mind that everything are operators rather than simple numbers) we have

$$\vec{L} = \vec{r} \times \vec{p} \Rightarrow r_l p_m \epsilon_{lmn} = L_n$$

$$[L_i r_j] = L_i r_j - r_j L_i = (r_l p_m \epsilon_{lmi} r_j - r_j r_l p_m \epsilon_{lmi})$$

$$= (r_l p_m r_j - r_j r_l P_m) \epsilon_{lmi}$$

since we have to sum offer l and m we can call $l = j$. indices that are summed over can be called ANYTHING you'd like and it is often convenient to call it another index which is also summed over. We call this "index gymnastics".

So

$$(r_j p_m r_j - r_j r_j p_m) \epsilon_{jmi}$$

$$= r_j [p_m r_j - r_j p_m] \epsilon_{jmi}$$

$$i\hbar \delta_{mj} r_j \epsilon_{jmi} = i\hbar \epsilon_{jmi} r_m = i\hbar \epsilon_{ijk} r_k$$

It takes some effort and practice but i hope you get the idea. Try your hand at deriving various classical vector identities in this manner (e.g. $a \times b \times c =$?. If not just write it ALL out with sums and all. by working many examples you'll become proficient and you'll assuredly find it's MUCH more handy than the whole vector notation!

So, we found our commutators this way (you derive the $\vec{p}$ commutator with $\vec{L}$. We find

$$[L_i, r_j] = i\hbar \epsilon_{ijk} r_k$$

$$[L_i, p_j] = i\hbar \epsilon_{ijk} p_k$$

The content of these expressions is to say " position and momentum transform like a vector" which is a sort-of big fat "duh", but we need these.

Now, we have

$$\vec{M} = \frac{1}{2m}(\vec{L} \times \vec{p} - \vec{p} \times \vec{L}) + \frac{km\vec{r}}{r}$$

Is this Hermetian? If so then it's obsservable. Your turn, find $[L_i, M_j]$

It is

$$[L_i, M_j] = i\hbar\epsilon_{ijk}M_k$$

And so it transforms as a vector just like we'd expect.

Now the BIG one

$$[M_i, M_j] = i\hbar\epsilon_{ijk}\left(-\frac{2H}{m}\right)L_k$$

We shan't derive this explicitly since it's a VERY nasty bit of algebra. It's AMAZING that the Hamiltonian pops out of the commutator almost by magic. I'm not familiar with any way we could've a priori predicted that! Nature is beautiful.

The derivation is aided by the observation that the commutator behaves analogously to the cross-product which we can write in index notation (there's something deep to this which we won't explore).

Viz,

$$[\vec{A}, \vec{B}]_k = A_i B_j \epsilon_{ijk}$$

then if

$$\vec{A} = \vec{r} \times \vec{p}$$

or

$$A_i = r_l p_m \epsilon_{lmi}$$

and

$$\vec{B} = \vec{r} \times \vec{p}$$

or

$$B_j = r_u p_v \epsilon_{uvj}$$

then we have, plugging in,

$$[\vec{A}, \vec{B}]_k = (r_l p_m \epsilon_{lmi})(r_u p_v \epsilon_{uvj})\varepsilon_{ijk}$$

$$= r_l p_m r_u p_v \epsilon_{lmi} \epsilon_{uvj} \epsilon_{ijk}$$

YIKES!!!

Let's think. Pondering we notice that all the terms with $r_l p_m \sim \delta_{lm}$ yield zero because δ_{lm} sends $\epsilon_{lmi} \to \epsilon_{lli} \equiv 0$

Similarly for $r_u p_v \to \delta_{uv} \to \epsilon_{uuj} \equiv 0$

The only option than doesn't is

$$p_m r_u \to i\hbar\delta_{mu}$$

Noticing this we may convert all the m's into u

$$[A, B]_k = r_l \left(i\hbar\delta_{mn}\right) p_v \epsilon_{lmi} \epsilon_{uvj} \epsilon_{ijk}$$

$$= i\hbar r_l p_v \epsilon_{lmi} \epsilon_{mvj} \epsilon_{ijk}$$

Now, we leverage an identity on the ϵ's.

$$\epsilon_{lmi} = -\epsilon_{mli}$$

by anti-symmetry and (a major identity)

$$\epsilon_{mli}\epsilon_{mvj} = \delta_{lv}\delta_{ij} - \delta_{lj}\delta_{iv}$$

Then

$$[A, B]_k = i\hbar r_l p_v \left[\delta_{lv}\delta_{ij} - \delta_{lj}\delta_{iv}\right] \epsilon_{ijk}$$

$$= i\hbar \left[r_v p_v \epsilon_{jjk} - r_j p_v \epsilon_{vjk}\right]$$

But $\epsilon_{jjk} \equiv 0$ by definition so that

$$[A, B]_k = i\hbar r_j p_v \epsilon_{vjk} = i\hbar r_i p_j \varepsilon_{ijk}$$

It's a mess, but if you're careful then the algebra is straight forward and kind of fun.

Now you try

$$[M_i, M_j] = -i\hbar\epsilon_{ijk}\frac{2H}{m}L_k$$

Next, for convenience later let's rescale M by

$$\mathscr{L}_i = \sqrt{\frac{-m}{2H}}M_i$$

Remember that H is an operator!
then

$$[\mathscr{L}_i, \mathscr{L}_j] = i\hbar\varepsilon_{ijk}L_k$$

Nice, tidy, and elegant.
All together our entire central potential algebra is then

$$\begin{aligned}
[L_i, L_j] &= i\hbar\varepsilon_{ijk}L_k \\
[M_i, L_j] &= i\hbar\varepsilon_{ijk}M_k \\
[M_i, M_j] &= -\frac{2i\hbar}{m}\varepsilon_{ijk}HL_k
\end{aligned}$$

This is our full symmetry generator algebra. From this everything else we'll study (states, etc) should follow.

Clearly $\vec{L}$ generates $SO(3)$. The algebra is closed. We don't get random stuff from it.

But $\vec{M}$ doesn't close, so it appears at first glance there's no further symmetry.

Maybe rescaled units will be clearer. Let's write everything in terms of $\mathscr{L}$

$$[\mathscr{L}_i, L_j] = i\hbar\epsilon_{ijk}\mathscr{L}_k$$

$$[\mathscr{L}_i, \mathscr{L}_j] = i\hbar\epsilon_{ijk}L_k$$

If $[\mathscr{L}_i, \mathscr{L}_j] \to \mathscr{L}_k$ then this would be $SO(3)$ also, but they don't.

Notice that $\mathscr{L}$ and L sort of mix in a similar way. Perhaps we'll get something cleaner if we add and subtract the two algebras in just the right way (this is a common thing to do but at our level it is not obvious given our lack of experience with Lie algebra!)

Let's try!
Define

$$\vec{Q} = \frac{1}{2}(\vec{L} + \vec{\mathscr{L}})$$

$$\vec{K} = \frac{1}{2}(\vec{L} - \vec{\mathscr{L}})$$

Then YOU can show that

$$\begin{aligned}
[Q_i, Q_j] &= i\hbar\epsilon_{ijk}Q_k \\
[K_i, K_j] &= i\hbar\epsilon_{ijk}K_k \\
[Q_i, K_j] &= 0
\end{aligned}$$

This is shocking!

This says we have two completely independent $SO(3)$ groups that is not angular momentum $SO(3)$. It's $\vec{Q}$ and $\vec{K}$ $SO(3)$. It's a combined transformation of $\vec{L}$ and $\vec{M}$ and H all together!

Now we know in quantum land we need unitarity and we've also learned that the $SO(3)$ algebra is entirely isomorphic to $SU(2)$ algebra then we will say that the Hydrogen atom manifestly has an $SU(2) \otimes SU(2)$ symmetry whose generators are labeled by (q, k) where

$$\vec{Q}^2|q,k\rangle = i\hbar q(q+1)|q,k\rangle$$

$$\vec{K}^2|q,k\rangle = i\hbar k(k+1)|q,k\rangle$$

where $q, k = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$.

Each $SU(2)$ has dimension $(2s+1)$ sop the total dimension of $SU(2) \otimes SU(2)$ is $(2q+1)(2k+1)$ for some (q, k)

It all works just like spin but here it is as though our states have two seperate versions of spin simultaneously.

For example, recall that spin $1 \Rightarrow m = -1, 0, 1$. Three individual states. The same thing is happening here but we have a pair together at the same time!

Now, in the theory of Lie Alegebras there is a concept called "Casimir Operators". These are operators in the algebra that commute with all other operators.

We've seen these before without naming them when we studied spin representations. There we found the operator L^2 commuted with all the generators. Here we expect Q^2 and K^2 to commute but since we have states with (q, k) in $SU(2) \otimes SU(2)$ we will study

$$Q^2 + K^2$$

and

$$Q^2 - K^2$$

Check that these expressions commute for all K_i and Q_i. The number of Casimir operators is called the "rank" of the Lie algebra. $SU(2)$ has Rank 1 and $SU(2) \otimes SU(2)$ has rank 2

Now, there is another constraint from classical physics given by $\vec{L} \cdot \vec{M} = 0$

But this is the same thing as $\vec{Q}^2 - \vec{K}^2$!

Which means all states must be such that

$$q = k$$

so the full symmetry is restricted to the "diagonal subspace of $SU(2) \otimes SU(2)$.

Okay, Now for the other Casimir. This one turns out to be especially interesting.

$$Q^2 + K^2$$

Well this $\frac{1}{2}\left(L^2 + \mathscr{L}^2\right)$

or

$$= \frac{1}{2}\left(L^2 - \frac{m}{2H}M^2\right)$$

but since $[L, H] = [M, H] = 0$ then whatever states we have are definitely going to be energy eigenstates as well. So we can replace H with the real valued energy E in our expression.

Therefore,

$$\left(Q^2 + K^2\right)|q, k\rangle = [q(q+1) + k(k+1)]\hbar^2|q, k\rangle$$

easy' but $q = k$

therefore

$$\hbar^2 2k(k+1)|k, k\rangle$$

On the other hand we have

$$\frac{1}{2}\left(L^2 - \frac{m}{2E}M^2\right)|k, k\rangle$$

$$\left(\frac{1}{2}\hbar^2\ell(\ell+1) - \frac{m}{2E}M^2\right)|k,k\rangle$$

But what is $M^2|k,k\rangle$ (let's call the coulombic factor α instead of k for clarity)

We evaluate M^2 from the classical expression and plug it in

$$M^2 = \alpha^2 + \frac{2E}{m}\left(\hbar^2 + \ell^2\right)$$

and so

$$\frac{1}{2}\left(L^2 - \frac{m}{2E}M^2\right) = -\frac{m\alpha^2}{4E} - \frac{1}{2}\hbar^2$$

but this should be equal to

$$\hbar^2 2k(k+1)$$

Now solving this for E we find the energy of states labeled by k!!

$$E = -\frac{m\alpha^2}{2\hbar^2(2k+1)^2}$$

finally call $n = 2k+1$ and we arrive at Balmer's famous spectral series!

$$E_n = -\frac{m\alpha^2}{2\hbar^2 n^2} = -\frac{me^4}{8\pi\varepsilon_0\hbar^2 n^2}$$

Wow!
for each n we have $2k+1$ states of the same energy
It's been a massive undertaking!!
Let's recap.

1. Hamilton told us to study

$$\vec{L} \times m\vec{a}$$

we alternatively studied orbits in velocity space. Either way we were led to suspect some hidden symmetry

2. we recalled that symmetry $\Rightarrow$ conservation of something

3. So we next found a constant vector classically

$$\vec{a} = \vec{p} \times \vec{L} - \frac{\alpha m \vec{r}}{r}$$

4. We promoted this to a Quantum operator as per Dirac's advice

$$\vec{M} = \frac{1}{2m}(\vec{p} \times \vec{L} - \vec{L} \times \vec{p}) - \frac{\alpha}{r}\vec{r}$$

5. We next checked if $\vec{M}$ was actually a symmetry generator by evaluating $[\vec{M}, \vec{H}] = 0$ since symmetries of H commute with H

6. Since it commuted with H then found all the relevant commutators in hopes of stumbling upon an algebra we recognized from previous study. If we could find no algebra we could always study the full theory of Lie Groups and find the one that fit. We know it would be there since we found this hidden symmetry.

7. We computed all the commutators to develop the algebra

$$\begin{aligned}
[L_i, L_j] &= i\hbar\epsilon_{ijk}L_k \\
[M_i, L_j] &= i\hbar\epsilon_{ijk}M_k \\
[M_i, M_j] &= -\frac{2i\hbar}{m}\epsilon_{ijk}HL_k
\end{aligned}$$

8. We didn't recognize anything other than $\vec{L}$'s algebra so we decided to rescale $\vec{M}$ in hopes of seeing something we could recognize

$$\mathscr{L}_i = \sqrt{\frac{-m}{2H}}M_i$$

9. so that we found

$$\begin{aligned}
[\mathscr{L}_i, L_j] &= i\hbar\epsilon_{ijk}\mathscr{L}_k \\
[\mathscr{L}_i, \mathscr{L}_j] &= i\hbar\epsilon_{ijk}L_k
\end{aligned}$$

10. Next we noticed this is tantalizingly close to a pair of $SO(3)$ or equivalently $SU(2)$ algebras. This prompted us to look at

$$Q = \frac{1}{2}(\vec{L} + \vec{\mathscr{L}})$$
$$K = \frac{1}{2}(\vec{L} - \vec{\mathscr{L}})$$

11. This surprisingly led to a pair of $SO(3)$ algebras.

$$[Q_i, Q_j] = i\hbar\epsilon_{ijk}Q_k$$
$$[K_i, K_j] = i\hbar\epsilon_{ijk}K_k$$

12. Since were in Quantum Land we identified $SU(2) \otimes SU(2)$ and then we studied the Casimir operators of this algebra

$$Q^2|q,k\rangle = \hbar^2 q(q+1)|q,k\rangle$$
$$K^2|q,k\rangle = \hbar^2 k(k+1)|qk\rangle$$

where $(q,k) = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$

13. Next we noticed classically that

$$\vec{L} \cdot \vec{M} = 0$$

But we found was $Q^2 - K^2 = \vec{L} \cdot \vec{M}$. This implied $q = k$ so that we may label all eigenstates by
k where for each value of k there are $(2k+1)$ states. for example, $k = \frac{1}{2} \Rightarrow (+1, -1)$ states just like with spin-$\frac{1}{2}$

14. Since we studied $Q^2 - K^2$ it made sense to look at $Q^2 + K^2$

we found

$$\begin{aligned} Q^2 + K^2 &= \frac{1}{2}\left(L^2 + \mathscr{L}^2\right) \\ &= \frac{1}{2}\left(L^2 - \frac{m}{2H}M^2\right) \end{aligned}$$

15. We noticed that since $[H, L] = 0 = [H, M]$ all of the eigenstates will be of the form $H|k\rangle = E_k|k\rangle$.

$\frac{1}{2}\left(L^2 - \frac{m}{2E}M^2\right)$ has eigenvalues

$$-\frac{1}{2}\hbar^2 - \frac{m}{4E}k^2$$

but on the other side we have

$$q^2 + k^2 = 2k(k+1)\hbar^2$$

so then

$$-\frac{1}{2}\hbar^2 - \frac{m}{4E}k^2 = 2k(k+1)\hbar^2$$

allowing us to solve for E_k

17. then

$$-\bar{E}_n = -\frac{mk^2}{2\hbar^2 n^2}$$

where $n = 2k + 1$

That's the solution of Hydrogen. It was long but definitely worth it.

It turns out that the diagonal subspace of $SU(2) \otimes SU(2)$ or $SO(3) \otimes SO(3)$ is $SO(4)$: the rotations in 4 dimensions.

The Hydrogen atom possesses $SO(4)$ symmetry. These most definitely arent (x, y, z, t) rotations but instead very abstract rotations within phase space involving conservation laws.

Finally we've arrived at Hydrogen.

States are labeled by the $SO(4)$ Lie Algebra.

We definitely have angular momentum states

$$L^2|\ell, m\rangle = \ell(\ell+1)|\ell, m\rangle$$
$$L_z|\ell, m\rangle = m_\ell(\ell, m\rangle$$

but now we also have the n labels. so our Hydrogen atom should have states labeled by

$$|n, \ell, m\rangle$$

and

$$\hat{H}|n, l, m\rangle = -\frac{m\alpha^2}{2\hbar^2}\frac{1}{n^2}|n, l, m\rangle$$

For each n there are $(2\ell+1)^2$ states of the same energy

$$\begin{aligned} n &= 1 \quad l = 0 \\ n &= 2 \quad l = 0, 1 \\ n &= 3 \quad l = 0, 1, 2 \end{aligned}$$

and so on.

What about states in the position basis? can we get the "traditional" electron orbitals from our algebraic method?

YES! we compute

$$\langle \vec{r}, t \mid n, l, m \rangle = \psi(\vec{r}, t)$$

Suppose we have the ground state

$$\mid 1, 0, 0 \rangle$$

viz, $n = 1$ and $\ell = m = 0$
then we can show that $\vec{Q}|1\rangle = 0$ and $\vec{K}|1\rangle = 0$
since $i, k = 0$
But

$$\vec{L} = \vec{Q} + \vec{K}$$

and

$$\vec{\mathscr{L}} = \vec{Q} - \vec{K}$$

but $\vec{M} \propto \vec{\mathscr{L}}$ so then $\vec{M}|1\rangle = 0$
but $\vec{M} = \frac{1}{2m}[\vec{p} \times \vec{L} - \vec{L} \times \vec{p}] - \frac{k\vec{r}}{r}$
so we operate on the ground state with $\vec{M}$

$$\vec{M}|1\rangle = \left[\frac{1}{2m}[\vec{p} \times \vec{L} - \vec{L} \times \vec{p}] - \frac{k\vec{r}}{r}\right]|1\rangle = 0$$

we can write this as (show this)

$$(\vec{p} \times \vec{L} - i\hbar\vec{p} - mk\hat{r})|1\rangle = 0$$

but $p = -i\hbar\nabla$ and we saw $\vec{p} \times \vec{L}|1\rangle = 0$
therefore

$$(i\hbar\vec{p} + mk\hat{r})|1\rangle = 0$$

gives us a first order differential equation for our position basis ground state!

$$\hbar^2 \frac{d}{dx}\psi(x) = -mk\psi(x)$$

$$\psi_{100}(r) = \left(\frac{1}{a_0}\right)^{3/2} 2e^{-r/a_c}$$

where $a_0 = \frac{\hbar^2}{me^2}$ is the Bohr radius which is the average distance of the electron from the proton in the hydrogen atom!

Spherical Coordinates

Recall polar coordinates: for systems with circular symmetry in two dimensions we often use polar coordinates instead of cartesian coordinates. This choice simplifies our description of rotationally symmetric systems drastically. For example, on the earth it's MUCH easier to say "I'm at latitude x and longitude y" instead of "from the center of the earth I am at (x,y,z)". Let's see how this works.

For polar coordinates we are familiar with the coordinate transformation

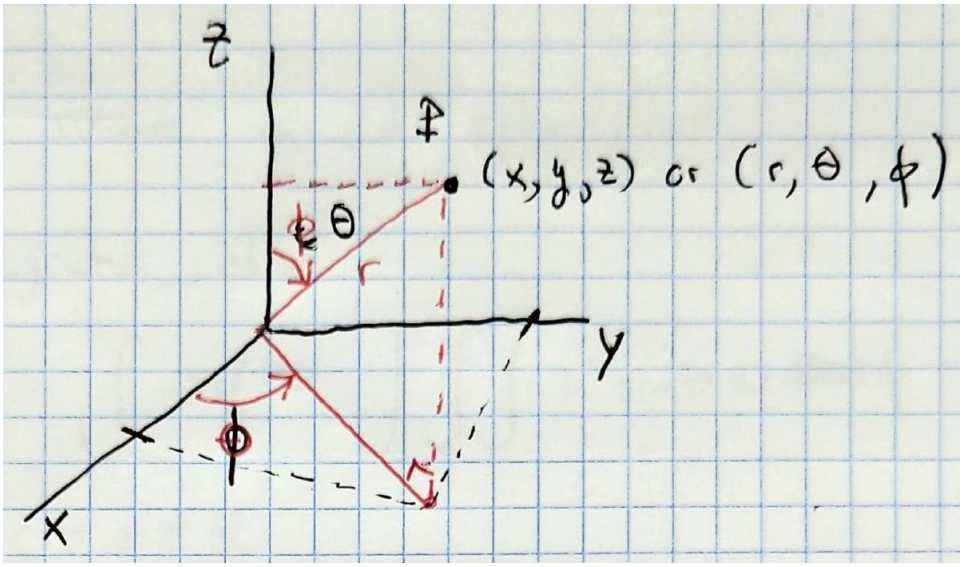
$$x = r\cos\theta \quad y = r\sin\theta$$

and their differentials

$$dx = \cos\theta dr - r\sin\theta d\theta$$

$$dy = \sin\theta dr + r\cos\theta d\theta$$

Similarly in three dimensions we sometimes use cylindrical or spherical coordinates if a system we are studying suits it. Later we shall be solving the Hydrogen atom which we expect should exhibit some degree of spherical symmetry. Therefore let us
establish spherical coordinates as in the diagram below



The Cartesian coordinates of P are expressed in spherical coordinates as

$$x = r\sin\theta\cos\phi$$
$$y = r\sin\theta\sin\phi$$
$$z = r\cos\theta$$

A volume element in 3D Cartesian coordinates is easily shown to be is $dV = dxdydz$. In spherical coordinates we find

$$dV = r^2\sin\theta d\theta d\phi dr$$

via either transformation of $dxdydz$ or by geometry.

For example, we may use this to find volume of a ball. We merely integrate dV over the entire volume in spherical coordinates:

$$V = \int dV = \iiint r^2\sin\theta dr d\theta d\phi$$

where

$$0 \leqslant r \leqslant R$$

$$0 \leq \theta \leq \pi$$

$$0 \leq \phi \leq 2\pi$$

(when we sweep around ϕ by 2π we cover half of θ. So then, we integrate dV

$$V = \int_0^R r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi = \frac{4}{3}\pi R^3$$

Now, in these coordinates we also wish to find a way to write derivatives in terms of our new spherical coordinates.

In Cartesian coordinates we have

$$\frac{\partial}{\partial x} \equiv \partial_x; \frac{\partial}{\partial y} \equiv \partial_y; \frac{\partial}{\partial z} \equiv \partial_z$$

and our spherical coordinates are in terms of (r, θ, ϕ):

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}$$

Then we have

$$dx = (\sin\theta\cos\phi)dr + (r\cos\theta\cos\phi)d\theta - r\sin\theta\sin\phi d\phi$$

$$dy = (\sin\theta\sin\phi)dr + (r\cos\theta\sin\phi)d\theta + r\sin\theta\cos\phi d\phi$$

$$dz = \cos\theta dr - r\sin\theta d\theta$$

Next we invert these so that we may find

$$\begin{aligned} dr &= \ldots \\ d\theta &= \ldots \\ d\phi &= \ldots \end{aligned}$$

Ultimately we seek the form of $\vec{\nabla}$ where $\vec{\nabla} = \begin{bmatrix} \partial_x \\ \partial y \\ \partial z \end{bmatrix}$

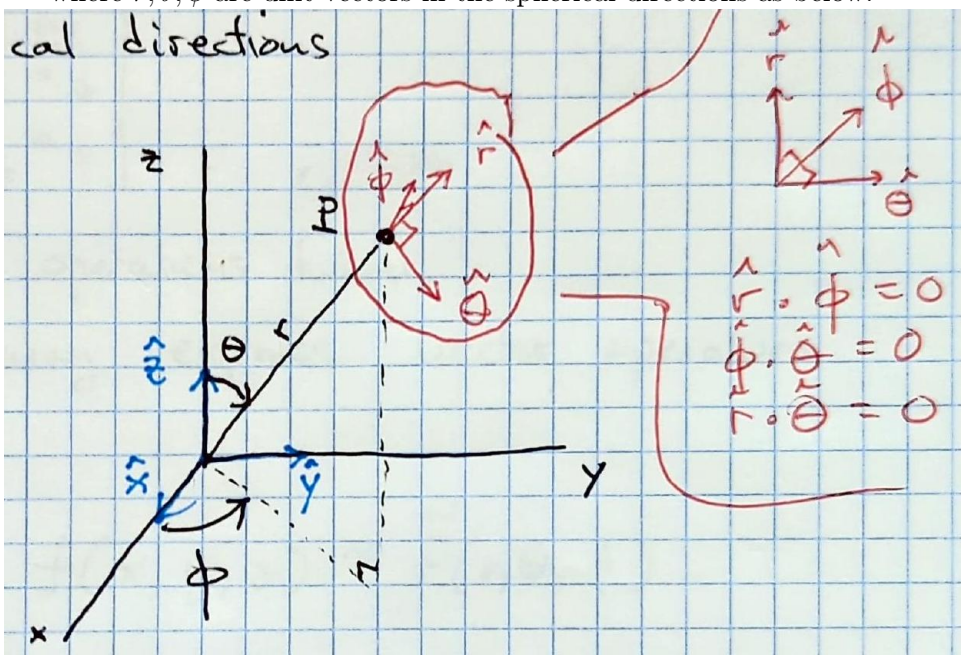
after lots of tedious algebra (do it!) we find

$$\nabla = \begin{bmatrix} \partial_r \\ \frac{1}{r}\partial_\theta \\ \frac{1}{r\sin\theta}\partial_\phi \end{bmatrix}$$

or

$$\nabla = \hat{r}\frac{\partial}{\partial r} + \frac{\hat{\theta}}{r}\frac{\partial}{\partial \theta} + \frac{\hat{\phi}}{r\sin\theta}\frac{\partial}{\partial \phi}$$

where $\hat{r}, \hat{\theta}, \hat{\phi}$ are unit vectors in the spherical directions as below.



We can also express cartesian unit vectors $\hat{x}, \hat{y}, \hat{z}$ in terms of the spherical unit vectors $\hat{r}, \hat{\theta}, \hat{\phi}$ as

$$\begin{aligned}
\hat{r} &= \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + \cos\theta\hat{z} \\
\hat{\theta} &= \cos\theta\cos\phi\hat{x} + \cos\theta\sin\phi\hat{y} - \sin\theta\hat{z} \\
\hat{\phi} &= -\sin\phi\hat{x} + \cos\phi\hat{y}
\end{aligned}$$

Such that a general position vector in three dimensions can be givben by

$$\vec{r} = r\sin\theta\cos\phi\hat{x} + r\sin\theta\sin\phi\hat{y} + r\cos\theta\hat{z}$$

In this system of spherical coordinates we may now write any arbitrary vector as

$$\begin{aligned}
\vec{a} &= a_r\hat{r} + a_\theta\hat{\theta} + a_\phi\hat{\phi} \\
&= \begin{bmatrix} a_r \\ a_\theta \\ a_\phi \end{bmatrix}
\end{aligned}$$

and then all vector operations follow.
Similarly we may express vector operators as:
for any function $f(x,y,z) \approx \tilde{f}(v,\theta,\phi)$

$$\nabla f = \frac{\partial f}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\hat{\phi}$$

$$\nabla \cdot \vec{a} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 a_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta a_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial a\phi}{\partial \phi}$$

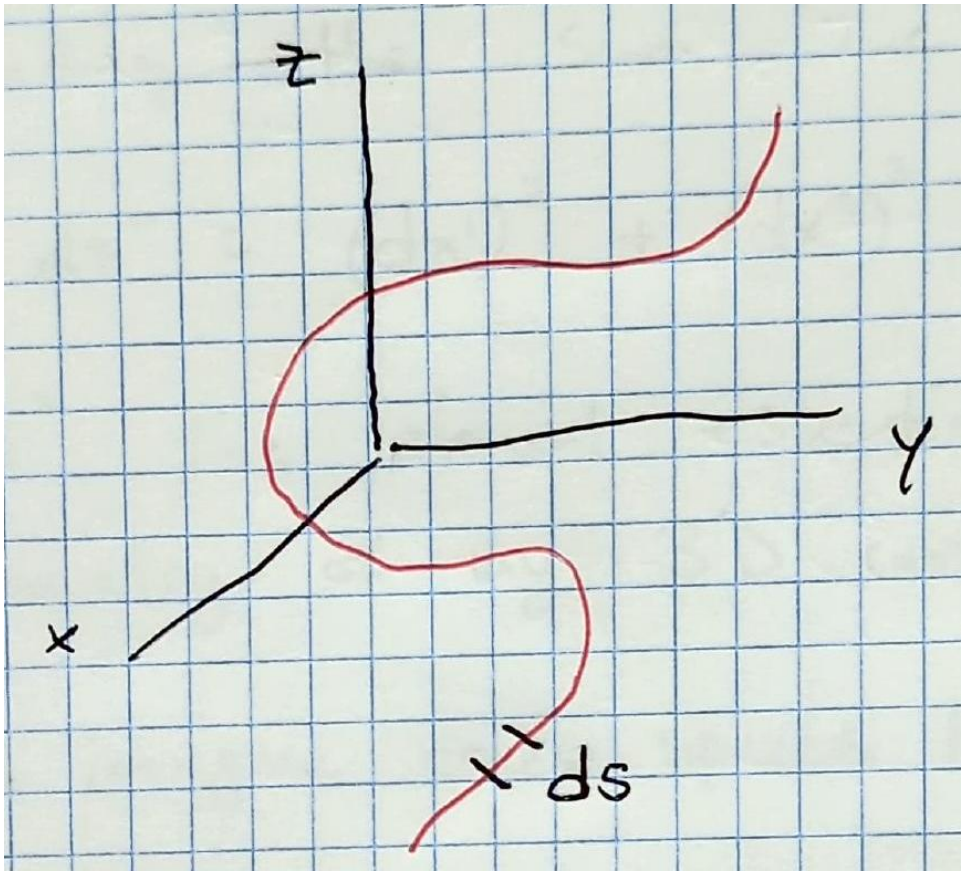
$$\nabla \times \hat{a} = \frac{1}{r^2\sin\theta}\det\begin{bmatrix} \hat{r} & r\hat{\theta} & r\sin\theta\hat{\phi} \\ \partial_r & \partial_\theta & \partial_\phi \\ a_r & ra_\theta & r\sin\theta a_\phi \end{bmatrix}$$

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2}$$

It would appear that we've made our lives absolutely horrible. However, as we will see later these coordinates simplify our lives greatly. For example, the Laplacian has a relatively simple solution in terms of spherical coordinates.

An aside on differential geometry

Consider the following:
suppose we have an arbitrary path in (Euclidian!) 3D space as below



We can define an incremental distance on this path as

$$ds^2 = dx^2 + dy^2 + dz^2$$

thanks to Pythagoras.

We could, if we really wanted, rewrite this as

$$ds^2 = \sum_{ij} dx^i dx^j \delta_{ij}$$

where $x^1 \equiv x \quad x^2 \equiv y \quad x^3 \equiv z$.

These superscripts only label the coordinate (a rose by any other name...). If I wanted to square x^2 I would write $\left(x^2\right)^2$ like so. (Yes yes, it feels really dumb to write things like this but it is very practical as we'll see).

$$\delta_{ij} = 1 \quad \text{if } i = j \quad ; \quad \delta_{ij} = 0 \quad \text{if } i \neq j$$

Expanding the sum we get back Pythagoras:

$$ds^2 = \left(dx^1\right)^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2$$

This ds^2 line element essentially defines the geometry of my 3D Cartesian and Euclidean space. We could imagine crazy spaces like

$$ds^2 = dxdy + (dz)^2 - xdydz$$

and then study the geometry through our "Pythagoras-like" line element. The Pythagorean element defines a Euclidean geometry. For a long time people couldn't imagine there could be anything else!

But our crazy example would be some strange geometry indeed! What on earth could parallel lines even be in that geometry?!

Well, it's always a good idea to generalize and then study the generalities rather than relying on specific examples. The most general geometry could be written as

$$ds^2 = g_{ij}dx^i dx^j$$

Here we have left out the $\sum_{ij}$ symbol and follow Einstein's rule:

"if an index repeats in the up and down position then we sum over that index from 1 to the dimension of our space D"

g_{ij} is called "the metric" and encodes the full geometry of our space.

For example, you may have heard that in special relativity (SR) space-time is "non-Euclidean". In SR we have

$$ds^2 = -c^2dt^2 + (dx)^2 + (dy)^2 + (dz)^2$$

$$= \eta_{\mu\nu}dx^\mu dx^\nu$$

where

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The ic difference between space and time coordinates makes ALL the confusing and beautiful difference between Newton and Einstein. This metric implies a non-Euclidean relationship between space and time coordinates.

Now, back to 3D Euclid land. Our Cartesian 3D world is entirely flat and obeys Pythagoras. So our line element is

$$ds^2 = \delta_{ij} dx^i dx^j$$

in (x, y, z) coordinates.

Now we convert to spherical coordinates (r, θ, ϕ) and find

$$ds^2 = g_{ij} dx^i dx^j$$

where

$$g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{bmatrix}$$

Importantly, we have NOT changed our space, we have ONLY changed how we describe the space. In our transformation we did not change in the slightest the fundamental geometry of the space. but this transformed geometry looks nothing like Pythagoras (or at least δ_{ij})!

You can, however, work out the details of the Pythagorean theorem and naturally find you get exactly the same results. So although the metrics "look" different they are, in a more general sense "the same". This is called "general coordinate invariance" - it simply means we are free to use ANY coordinates we like so long as the fundamental geometry doesnt change.

so a natural question to ask is "when is g_{ij} equivalent to g'_{ij}?" Two metrics may look entirely different (like our crazy example) but may represent the same geometry.

As an example the 2D plane is manifestly Pythagorean and Euclidean - parallel lines never intersect (except maybe at infinity).

If, however, we are an ant living on a spherical surface then our space is fundamentally NOT flat and non-Euclidean - parallel lines most assuredly intersect. In Polar coordinates the (x, y) plane becomes

$$ds^2 = dr^2 + r^2 d\theta^2$$

However, for a spherical surface we have

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2$$

and these two 2D spaces are manifestly and fundamentally different. There is no equivalence between

$$g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}$$

and

$$g'_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}$$

This is the subject of differential geometry and we need not say any more. This is the very beginning of the study of one of the most beautiful areas of mathematics that exists and extreme importance in modern physics.

Our main goal is getting a handle on spherical coordinates in Euclidean flat space - no need yet for non-Euclidean geometries!

Hydrogen

Now that we've got some tools under our belt it's time to try and model the hydrogen atom.

Classically it is modeled as a two particle bound state of a massive proton of positive charge and a very light electron of negative charge.

The problem is, however, that in the classical solution Maxwell says "accelerating charges radiate electro-magnetic energy" and Newton says "if you have uniform circular motion then there is an inwardly point acceleration - the centripetal acceleration"

By considering these two facts it can be shown that an electron in orbit about a proton would spiral into the proton VERY quickly; much too quickly to match observations of stable hydrogen!

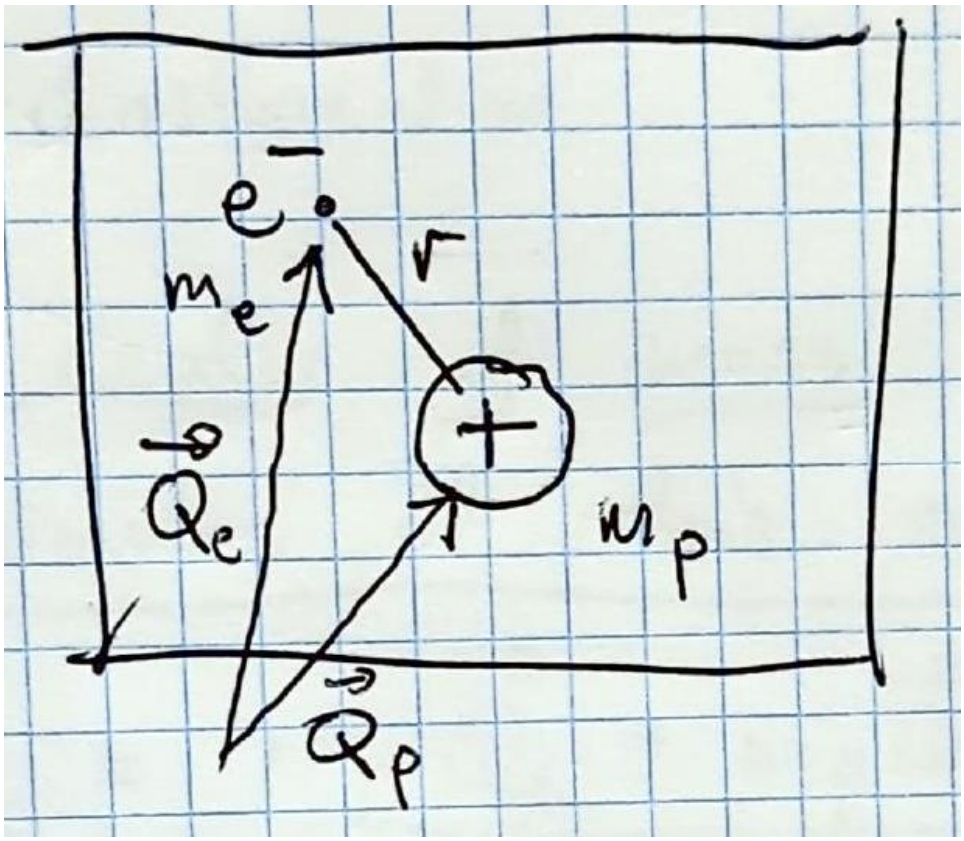
So clearly something is wrong with our description. Quantum mechanically, Heisenberg says that "if you know the position of a quantum object then you know nothing of it's momentum" - (uncertainty relations). Clearly if the electron were to spiral down into the proton then we would know the position of the electron extremely well but this cant happen. On top of this fuzzy explanation is from de Broglie who says "all matter has an associated wavelength" and the wavelength of an electron in an atom is much bigger than the size of the nucleus.

Clearly we seek a quantum mechanical description of this system but how? The simplest and most naïve approach is to write down the classical model, promote position and momentum into operators, and simply hope for the best.

Shockingly, this approach turns out being VERY accurate! There are some subtleties that we end up missing (Lamb shift, hyperfine splitting, etc) but this procedure matches what we know about hydrogen VERY well. We

can always add in bells and whistles to our model later if needed (relativity, spin, quantum field theory, QED, etc). Ancient proverb - "make a problem as simple as possible without being trivial".

Therefore, we shall model hydrogen as a two-particle system composed of a heavy positive proton and a very light negative electron. Classically we imagine the bound system as below



We can write down our classical Hamiltonian as

$$H = \frac{\vec{p}_e^2}{2m_e} + \frac{\vec{p}_p^2}{2m_p} + V(\vec{r})$$

where m_e is the electron mass, m_e is the proton mass, and $V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}$ is the Coulomb potential.

Let us add a coordinate system and say that e^- is at position $\vec{Q}_e$ and p^+ is at $\vec{Q}_p \Rightarrow r = \left|\vec{Q}_e - \vec{Q}_p\right|$

For simplicity we define units such that $\alpha = \frac{1}{4\pi\varepsilon_0} \equiv 1$
So then

$$H = \frac{\vec{p}_e^2}{2m_e} + \frac{\vec{p}_p^2}{2m_p} - \frac{e^2}{\left|\vec{Q}_e - \vec{Q}_p\right|}$$

This is exact (insofar as our model goes) but now we make a coordinate transformation.

Consider the center of mass and relative coordinates of this system similar to what we've done in our conservation of momentum studies. We say "i don't care about e^- or p^+ per se, i care about their relative orientation to each other". So we have

$$\vec{Q}_r = \frac{m_e\vec{Q}_e + m_p\vec{Q}_p}{m_e + m_p}$$

$$\vec{Q}_c = \vec{Q}_e - \vec{Q}_p$$

and

$$\vec{P}_c = \vec{P}_e + \vec{P}_p$$
$$\vec{P}_r = \frac{m_p\vec{P}_e - m_e\vec{P}_p}{m_e + m_p}$$

This is called a "canonical transformation" in classical physics.

Now, we may rewrite the Hamiltonian in these new coordinates

$$H = \frac{P_c^2}{2\left(m_e + m_p\right)} + \frac{P_r^2}{2\mu} - \frac{e^2}{\left|\vec{Q}_r\right|}$$

where $\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_p}$ is called the "reduced mass" of the system.

Notice what happened: the center of mess behaves like a free particle of mass $M_H = m_e + m_p$: like a complete hydrogen atom. Hydrogen as a bound system can move about with some momentum. I can throw hydrogen at you so certainly i should be able to find a total momentum or kinetic energy term for the full system.

But look! The interaction term says nothing about the center of mass! If we want we can study this system such that the COM is fixed: i.e. we have a static hydrogen atom with zero kinetic energy.

So we are free to take $P_c = 0$ We can always add it back in using a Galilean or Lorentz boost.

So then, the interesting physics we wish to pursue here is given by

$$H = \frac{P_r^2}{2\mu} - \frac{e^2}{\left|\vec{Q}_r\right|}$$

This is what we shall solve.

Now, let's drop the subscript r and follow our canonical quantum procedure Dirac taught us. We put operator hats on P and Q

$$\hat{H} = \frac{\hat{\vec{P}}^2}{2\mu} - \frac{e^2}{|\hat{\vec{Q}}|}$$

where $\vec{P}$ and $\vec{Q}$ have three components each

$$\hat{P}_x, \hat{P}_y, \hat{P}_z \quad : \quad \hat{Q}_x, \hat{Q}_y, \hat{Q}_z$$

Our next step is to find the energy eigenstates of this Hamiltonian/system. This is not trivial and we will have to work hard.

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

Now, we shall work in the position basis and therefore will study

$$\langle x, y, z|\psi\rangle = \psi(x, y, z)$$

We anticipate spherically symmetric solutions for ψ

Now, in position space we replace we have

$$\hat{\vec{P}} = -i\hbar\vec{\nabla}$$

$$\hat{\vec{Q}} = \vec{Q}$$

for example $P_x = -i\hbar\frac{\partial}{\partial x}$ and similarly for y, z

then

$$H\psi = E\psi$$

Plugging in our operators we have

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi(\vec{r}) - \frac{e^2}{r}\psi(\vec{r}) = E\psi(\vec{r})$$

we wish to find all solutions to to this exceedingly complicated looking second order non-homogeneous partial differential equation!

Anticipating spherically symmetric solutions we choose spherical coordinates in hopes that we find some simplicity.

Some general advice when faced with really complicated differential equations.

1. check the large/small variable limits. Here we expect free particle eigenstates for large r. Indeed taking this limit produces a free particle as expected: the potential goes to zero.

2. try something. Here we will consider solutions that are "separable". That is to say, of the form

$$\psi(\vec{r}) = R(r)\Theta(\theta)\Phi(\phi)$$

This is what we mean by "separable". Not all functions are separable (e.g $f(x,y) = x^2 + xy \neq g(x)h(y)$) and so this is not guaranteed to solve our equation. But it is generally a good first thing to try in order to get a foothold on the problem and it turns out to work just fine!

3. Now we just have to write ∇^2 in spherical coordinates.

$$\nabla^2 \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial}{\partial r}\right] + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\right] + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial\phi^2}$$

Again this form of ∇^2 looks like a nightmare but we roll with it.

For ease let's separate ψ into radial and angular parts and plug into our equation (we will worry about whether $Y = \Theta\Phi$ separates into pieces later).

$$\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$$

Then, plugging in, we have

$$-\frac{\hbar^2}{2\mu}\left[\frac{Y}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{R}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{R}{r^2\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right] - \frac{e^2}{r}RY = ERY$$

Now, we divide by RY and multiply by $-\frac{2\mu r^2}{\hbar}$ then

$$\frac{1}{R}\frac{d}{dr}\left[r^2\frac{dR}{dr}\right] - \frac{2\mu r^2}{\hbar^2}[V(r) - E]$$
$$+\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right] = 0$$

Now dwell upon what this says.

We have a differential equation for $R(r)$ alone plus a differential equation for $Y(\theta, \phi)$ and they happen to add to produce zero!

The only way to get zero is if the $R(r)$ equation equals some constant and the $Y(\theta, \phi)$ equals minus that constant. So we have $a - a = 0$, duh!

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2\mu r^2}{\hbar^2}\left(-\frac{e^2}{r} - E\right) = a$$

and

$$\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 y}{\partial\phi^2}\right] = -a$$

But look! We have two separate differential equations. We may solve these separately!

Let's start by solving for $Y(\theta, \phi)$

Multiply through by $Y\sin^2\theta$ to get

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{\partial^2 Y}{\partial\phi^2} = -Ya\sin^2\theta$$

Now we try separating $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$. We plug this into our differential equation for Y and divide by $\Theta\Phi$ to get

$$\frac{1}{\Theta}\left[\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)\right] + \frac{a}{\Theta}\sin^2\theta + \frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2} = 0$$

Now we utilize the same trick again! We have two separate differential equations that sum to zero!

the Θ equation must equal a constant m^2 and the Φ equation is $-m^2$ because

$$m^2 - m^2 = 0$$

Now, to be fair there's absolutely no reason to write m^2 rather than m or anything else. We only call it m^2 because in hindsight it makes life easier.

So then

$$\frac{1}{\Theta}\left[\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\theta}{d\theta}\right)\right]+a\sin^2\theta=m^2$$

and

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2}=-m^2$$

Dear reader, by now you're a pro at the Φ equation!

$$\frac{d^2\Phi}{d\phi^2}=-m^2\Phi\Rightarrow\Phi(\phi)=e^{\pm im\phi}$$

modulo constants which we will put in later. The functional form is most important here.

Furthermore we demand that

$$\begin{aligned}\Phi(\phi+2\pi n)&=\Phi(\phi)\\ &\Rightarrow m=0,\pm1,\pm2,\ldots\end{aligned}$$

So we got one of the three functions. Unfortunately the other two are definitely not as straightforward!

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\left(a\sin^2\theta-m^2\right)\Theta=0$$

We shall not derive the full solution of this differential equation but it is a standard and common set of functions that occur all over the place in mathematics. Here we cite the result,

$$\Theta(\theta)=AP_\ell^m(\cos\theta)$$

where

$$P_\ell^m(x)=\left(1-x^2\right)^{|m|/2}\left(\frac{d}{dx}\right)^{|m|}P_\ell(x)$$

and

$$P_\ell(x)=\frac{1}{2^\ell\ell!}\left(\frac{d}{dx}\right)^\ell\left(x^2-1\right)^\ell$$

Yes, yes...the first time we're exposed to Legendre functions we feel like idiots and maybe feel a little discouraged by their ugly form. As you get

to know them, however, their beauty will slowly fill your soul as they have numerous students in history!

The ℓ comes from our choice of constant a

Here we let $a = -\ell(\ell+1)$ and use ℓ from now on (but you already have an idea what this ℓ is!)

$P_\ell^m(X)$ are called "the associated Legendre polynomials"

$P_\ell(x)$ are called" the ℓ^{th} Legendre polynomial"

The first few are

$$\begin{aligned} P_0 &= 1 \\ P_1 &= x \\ P_2 &= \frac{1}{2}\left(3x^2 - 1\right) \end{aligned}$$

and so on

Here $P_\ell^m(\cos\theta)$ is always a polynomial in $\cos\theta$ (or $\sin(\theta)$ via phase shift).

$$\begin{aligned} &P_0^0 = 1 \quad P_1^1 = \sin\theta \quad P_1^0 = \cos\theta \\ &P_2^0 = \frac{1}{2}\left(3\cos^2\theta - 1\right) \end{aligned}$$

and so on. They're not excessively complicated.

Now, ℓ, m are labels for each function.

For any give value of ℓ there will be $(2\ell+1)$ values of m. what you've previously studied should be thoroughly screaming at you at this point!

possible values of ℓ, m are

$\frac{l}{0}$	$\frac{m}{0}$
1	$-1, 0, 1$
2	$-2, -1, 0, 1, 2$
3	$-3, -2, -1, 0, 1, 2, 3$

and so on!

These are angular momentum representations. ℓ corresponds to the irreducible representation of $SO(3)$; m corresponds to the eigenstates of that irrep! We secretly solved a bunch of complicated differential equations entirely algebraically following Lie and his symmetry group considerations! viz

$$\hat{L}_z|\ell, m\rangle = \hbar m|\ell, m\rangle$$

So we have $Y_\ell^m(\theta, \phi) = \Theta(\theta)\Phi(\phi)$

We haven't solved the radial equation yet but we can go ahead and normalize our $\psi(\vec{r})$. We require

$$\int |\psi(\vec{r})|^2 dV = 1$$

in spherical coordinates $dV = r^2 \sin\theta dr d\theta d\phi$
then

$$\int |\Psi|^2 r^2 \sin\theta dr d\theta d\phi = \int |R|^2 r^2 d\sigma \int |Y|^2 \sin\theta d\theta d\phi$$

we take $\int |R|^2 r^2 dr = 1$
and

$$\int |Y_\ell^m|^2 \sin\theta d\theta d\phi = 1$$

then, finally

$$Y_\ell^m(\theta,\phi) = \varepsilon \sqrt{\frac{(2\ell+1)(\ell-|m|)!}{4\pi(\ell+|m|)!}} e^{im\phi} P_\ell^m(\cos\theta)$$
$$\varepsilon = (-1)^m \quad \text{if } m \geq 0 \quad ; \quad \varepsilon = 1 \text{ for } m < 0$$

Interestingly and importantly, all of these $Y_\ell^m(\theta,\phi)$ are orthogonal. That is to say,

$$\int_0^{2\pi} \int_0^{\pi} [Y_\ell^m(\theta,\phi)]^* [Y_{\ell'}^{m'}(\theta,\phi)] \sin\theta d\theta d\phi$$
$$= \delta_{\ell\ell'}\delta_{mm'}$$

This is just like in our Lie group approach where we said

$$\langle \ell, m \mid \ell', m' \rangle = \delta_{\ell\ell'}\delta_{mm'}$$

There are many other useful identities involving $Y_e^m(\theta,\phi)$ which we will not pursue here.

Additionally, what we ultimately found was

$$\langle \theta, \phi \mid \ell, m \rangle = Y_\ell^m(\theta,\phi)$$

$Y_\ell^m(\theta,\phi)$ are called "spherical harmonics". They intuitively describe oscillations on a sphere.

Now, all we have left is the lone radial equation.

$$\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right) - \frac{2\mu r^2}{\hbar^2}\left[-\frac{e^2}{r} - E\right] R = \ell(\ell+1)R$$

let $u(r) = rR(r)$
then

$$-\frac{\hbar^2}{2\mu}\frac{d^2u}{dr^2} + \left[V + \frac{\hbar^2}{2\mu}\frac{\ell(\ell+1)}{r^2}\right]u = Eu$$

This looks just like a one dimensional Schrodinger equation, with effective potential

$$V_{eff} = V + \frac{\hbar^2}{2\mu}\frac{\ell(\ell+1)}{r^2}$$

now, take $\kappa = \frac{\sqrt{-2mE}}{\hbar}$ and $E < 0$ for bound states.
We have

$$\frac{1}{\kappa^2}\frac{d^2u}{dr^2} = \left[1 - \frac{\mu e^2}{2\pi\epsilon_0\hbar^2\kappa^2 r} + \frac{\ell(\ell+1)}{(\kappa r)^2}\right]u$$

Now, yet again, set $\rho = \kappa r \quad \rho_0 = \frac{\mu e^2}{2\pi\varepsilon_0\hbar^2\kappa}$
then

$$\frac{d^2u}{d\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{\ell(\ell+1)}{\rho^2}\right]u$$

as $\rho \to \infty$ our equation and solution are simple

$$\frac{d^2u}{d\rho^2} = u \Rightarrow u(\rho) = Ae^{-\rho} + Be^{\rho}$$

we throw out e^{ρ} since this blows up as $\rho \to \infty$
Now as $\rho \to 0$

$$\frac{d^2u}{d\rho^2} = \frac{l(l+1)}{\rho^2}u$$

In this regime the centrifugal term completely dominates and we have

$$\Rightarrow \quad u(\rho) = C\rho^{\ell+1} + D\rho^{-\ell}$$

or, since we're near $\rho = 0$

$$u(\rho) \sim Ce^{\ell+1}$$

Therefore we expect our total solution to behave as

$$u(\rho) = \rho^{\ell+1}e^{-\rho}v(\rho)$$

Where we know nothing about $v(\rho)$ except it should be polynomial.

So we could proceed to pursue power series solutions of this differential equation by plugging in

$$v(\rho) = \sum_{j=0}^{\infty} c_j \rho^j$$

into

$$u(\rho) = \rho^{\ell+1} e^{-rho} v(\rho)$$

and then plugging all of THAT into our differential equation to try and find some kind of recursion relation between
c_{j+1} and c_j and doing so we would ultimately find that there is a maximum possible j allowed such that

$$2\left(j_{\max} + \ell + 1\right) - \rho_0 = 0$$

we call $j_{\max} + \ell + 1 = n$ where n is called the "principal quantum number". This allows us to determine the energy of the system as

$$E_n = -\left[\frac{\mu}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\right]\frac{1}{n^2}$$

where $n \in \mathbb{N}$

The radial functions then depend on both n and ℓ.

For fixed n, ℓ there exist $(2\ell + 1)$ states at that energy $m = -\ell, -\ell + 1, \ldots + \ell$

We have found finally,

$$\psi_{nlm}(r, \theta, \phi) = R_{n\ell}(r) Y_\ell^m(\theta, \phi)$$

where

$$v(\rho) = L_{n-\ell-1}^{2\ell+1}(2\rho)$$

where

$$L_{q-p}^p(x) = (-1)^p \left(\frac{d}{dx}\right)^p L_q(x)$$

$$L_q(x) = e^x \left(\frac{d}{dx}\right)^q \left(e^{-x} x^q\right)$$

$L_q(x)$ are called the Laguerre Polynomials
Finally the full hydrogenic wavefunction is

$$\psi_{nlm} = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-\ell-1)!}{2n[(n+1)!]^3}} e^{-\frac{r}{na}} \left(\frac{2r}{na}\right)^l \left[L_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{na}\right)\right] Y_\ell^m(\theta,\phi)$$

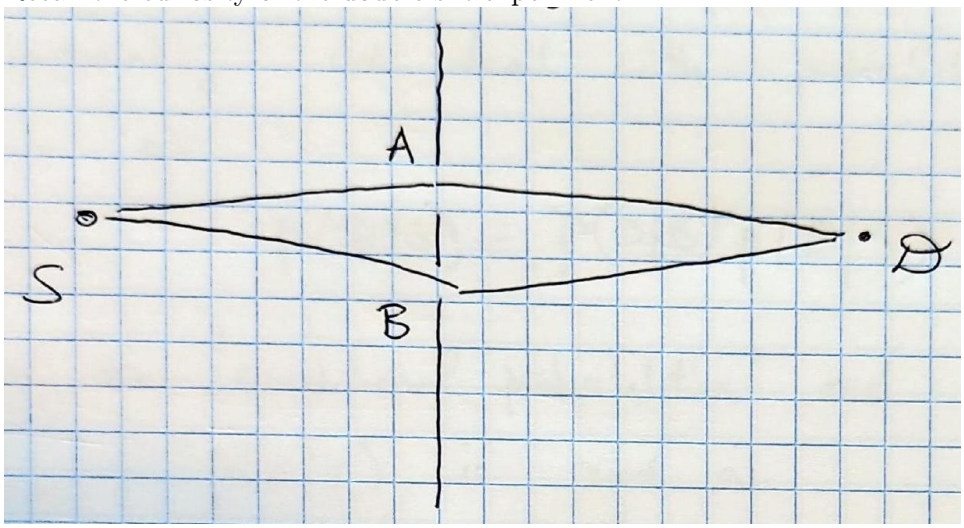
Where a is called the Bohr radius.

Now, caution: people often think of this as the "wavefunction of the electron" in hydrogen. However, we derived this as a relative coordinate system. The reason we can say "this is the electron's orbitals" is because since the proton is so darned heavy compared to the electron then the reduced mass is approximately equal to the proton's mass. For all practical purposes, we can get away with saying the is the wavefunction of the electron, but for very sensitive considerations we may have to remember that we chose center of mass coordinates.

Furthermore, this is a static hydrogen atom. if it were moving with some momentum P then this function would be multiplied by the translation operator e^{iPx} where P is the COM momentum.

Barandes' Stochastic - Quantum Correspondence

Recall the curiosity of the double-slit experiment



The amplitude for detection at D is given by,
amplitude of the A path + amplitude of the B path.

$$A(S \to D) = z_A + z_B$$

then the probabity of detection at $\mathcal{D}$ is

$$\begin{aligned} A^2 = |z_A + z_B|^2 &= \left(z_A^* + z_B^*\right)\left(z_A + z_B\right) \\ &= z_A^2 + z_B^2 + \left(z_A^* z_B + z_B^* z_A\right) \\ &= P_A + P_B + (\text{ interference }) \end{aligned}$$

The interference term $z_A^* z_B + z_A z_B^*$ is what makes Quantum Theory so different from Classical theory.

Classically we deal with probability rules such as

$$P(D \mid S) = P(D \mid B)p(B \mid S) + p(D \mid A)_P(A \mid S)$$

these are "conditional probabilities" and $p(x \mid y)$ is read as
"the probability of observing x given you know y"
"
For example, the probability it will rain today given that it's cloudy would be greater than the probability it will rain given that it's sunny.

$$P(r \mid c) \geq P(r \mid s)$$

This is, in general, different than the straight probability it will rain, evidently

$$P(\text{ rain }) = P(r \mid c) + P\left(r \mid s\right)$$

and $\quad P(\text{ rain }) + P(\text{ no rain }) = 1$

In general given a set of possibilities (called a sample space)
we have $\sum_i P(i) = 1$

$$P(i) = \sum_i P\left(i|j\right) P(j)$$

This is called "Bayesian Marginalization" and an important cornerstone of probability theory.

Additionally we also have Bayes' Theorem

$$P(A \mid B) = \frac{P(A, B)}{P(B)}$$

where $P(A, B)$ is the joint probability of BOTH A and B occurring

For example, the probability that it rains AND is cloudy

$$P(\text{ rain } \mid \text{ cloud }) = \frac{P(\text{ rain AND cloudy })}{P(\text{cloudy})}$$

and so on.

These particular formulas are standard probability theory. For continuous spaces we have, similarly,

$$\int_{-\infty}^{\infty} p(x)dx = 1 \quad p(y) = \int p(y \mid x)p(x)dx$$

Where $p(x)$ is the "probability density function" (pdf).

So far these formula have been for random variables $\mathbb{X}$ and $\mathbb{Y}$

Now we imaging a random variable $\mathbb{X}$ that depends on time. $\mathbb{X}(t)$ is called a stochastic process. It could be discrete or continuous in time. Because $\mathbb{X}$ is a random variable it will have an associated probability density which also depends on time t;

$$p : \mathcal{C} \times \mathcal{T} \to [0, 1] \in \mathbb{R}$$

$\mathcal{C}$ is called the "configuration space" or "state space" and $\mathcal{T}$ is the set of times over which the process occurs. For example, the configuration space $\mathcal{C}$ of a dice is the set $[1, 2, 3, 4, 5, 6]$ and we can roll a dice every second thereby creating a (relatively uninteresting) stochastic process. It's relatively uninteresting because the probability at each time doesn't change. A slightly more interesting stochastic process would be rolling ever larger dice each timestep (i.e. coin, then 4 sided dice, then 6 sided, and on and on).

We shall have p be labeled by $i \in \mathcal{C}$ and $t \in \mathcal{T}$
by

$$p_i(t) \equiv p(i, t)$$

and shall consider mostly discrete configuration spaces. Continuous configuration spaces will be written as $p(x, t)$.

For example, stock market prices can be modeled by a stochastic process yielding a probability density for a price being between x and $x + dx$ dollars. You may naturally expect this probability distribution to change as a function of time, e.g. the price of apple stock a decade ago had a MUCH different probability distribution as it does now! Similarly the probability distribution of the location of ink dropped into water initially has a very narrow distribution which widens as a function of time until there is a probability of an ink molecule being found anywhere.

Now, $p_i(t)$ is the time dependent stand alone probability for configuration i at time t. We will typically set the initial probabilities of each configuration by hand. These are our initial conditions (e.g. the ink molecule is 100%

localized at $x = 0$ at $t = 0$ and then $p(x,t)$ evolves according to some equation in space and time.) viz,

$$\sum_j^N p_j(t=0) = 1$$

will be freely adjustable.

We will now call the conditional probability $p(i,t \mid j,0) = \Gamma_{ij}(t)$. This is also called a "transition probability of configuration $j \to i$.

We read this as "the probability of i at time t given that we have configuration j at $t = 0$."

For all times we demand (definition of probability) that

$$\sum_i p_i(t) = 1$$

The formula

$$p_i(t) = \sum \Gamma_{ij}(t) p_j(0)$$

will be the major component of our study. For continuous configurations this formula is or

$$p(x,t) = \int p(x,t \mid x',t') p(x',t') dx'$$

This formula tells us how our stochastic system evolves in space and time. In particular it looks like a matrix equation. viz,

$$\vec{p}(t) = \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_N(t) \end{bmatrix} \qquad \Gamma(t) = \begin{bmatrix} \Gamma_{11}(t) & \Gamma_{12}(t) & \dots \\ \vdots & \ddots & \\ \Gamma_{N1}(t) & \dots & \dots \end{bmatrix}$$

$\Gamma(t)$ is called a "Stochastic Matrix" if

$$\sum_i \Gamma_{ij}(t) = 1$$

i.e. the columns of $\Gamma(t)$ individually add to 1 and each $\Gamma_{ij}(t) \geq 0$

In particular $\Gamma(0) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbb{1}$ for a three state system.

Markov and non-Markov

Stochastic Processes can be broken up into two type.

1. Markov

Traditionally a Markov process is a stochastic process such that

$$p\left(x,t \mid x_m,t_m;x_{m-1},t_{m-1};\ldots;x_1,t_1\right)=p\left(x,t \mid x_m,t_m\right)$$

This natural seeming condition means that the conditional probability at t ONLY depends on the data at (x_m,t_m) (i.e. the previous value and NOT the total history).

This is often referred to as memoryless. Most systems of interest are assuredly not Markovian. For example, the price of apple stock in December strongly depends on the price the previous December, rather than just November, etc. Or position in the future depend on the velocity of an object which depends on previous positions.

2. We will consider an alternative form of Markovianity as

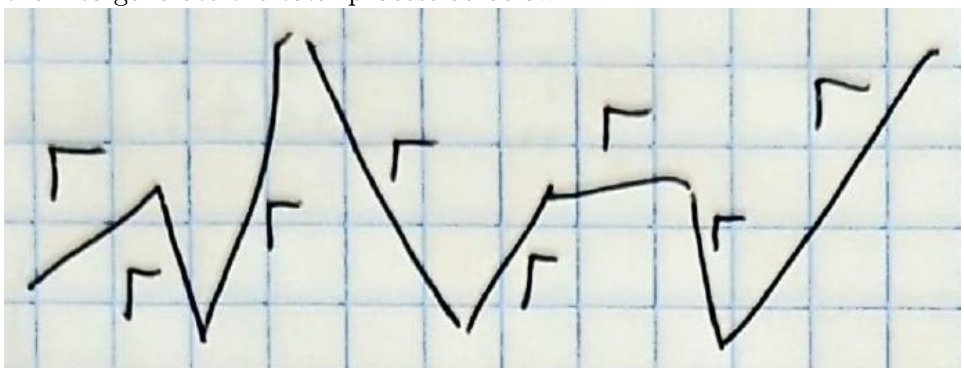
$$\Gamma_{ij}(t)=\sum_{k=1}^{N}\Gamma_{ik}\left(t''\right)\Gamma_{kj}\left(t'\right)$$

or simply

$$\Gamma(t)=\Gamma\left(t\leftarrow t'\right)\Gamma\left(t'\right)$$

This property is called "divisibility". Most physical systems are divisible. It's very difficult to imagine a physical process which would not satisfy divisibility. Divisibility is like playing the frames of a movie in succession.

This means we can chop our process up into small chunks and combine them to generate the total process as below.



If, for some reason, we cant do this we will say the process is "non-markovian" and "indivisible".

An example of an indivisible process would be

$$\Gamma(t) = \begin{bmatrix} \cos^2 \omega t & \sin^2 \omega t \\ \sin^2 \omega t & \cos^2 \omega t \end{bmatrix}$$

There does not exist any way that we can write this matrix as a product

$$\Gamma(t) = \Gamma\left(t \leftarrow t'\right) \Gamma\left(t'\right)$$

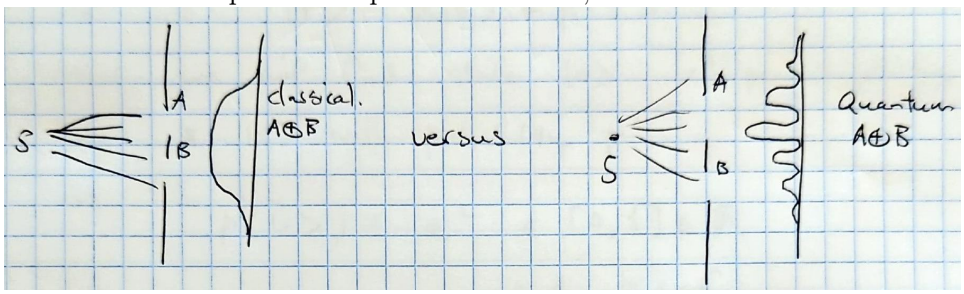
even though this is a perfectly reasonable stochastic matrix (in fact doubly stochastic!), viz, $\sin^2 \omega t + \cos^2 \omega t = 1$. What this process means, physically, i cannot say, but mathematically it is a perfectly valid stochastic process which happens to be indivisible.

Quantum Mechanics is NOT divisible: The Double Slit

Recall our discussion of the double slit. For a classical particle we expect the probabilities to be divisible. However, for a quantum object we have

$$\begin{aligned} p(\text{ detector } | \text{ source }) \neq\ & p(A \mid \text{ source }) p(\text{ detector } | A) \\ & + p(B \mid \text{ source }) p(\text{ detector } \mid B) \end{aligned}$$

If this were true then we wouldn't see an interference pattern at D. We would see the simple sum of probabilities. viz, below



Well in our quantum theory our probabilities arent divisible but we did sub-divide the system's Quantum Propagator when we derived the Feynman Path Integral. We claimed we could chop the amplitudes up into time steps and at the very end sqaure the total amplitude to get the probabilities, interference and all.

We have a question answer: **Does stochastic indivisibility $\Rightarrow$ quantum propagator divisibility?**

Surprisingly, it does! We will see that the difference between indivisible stochastic processes and divisible processes is precisely the interference part of the Quantum system! We will find

$$\Gamma(t) - \Gamma\left(t \leftarrow t'\right)\Gamma\left(t'\right) = \text{ Interference}$$

The total process is given by $\Gamma(t)$ and we subtract off the divisible part and obtain exactly the quantum interference.

This seems to hint that Quantum theory is a different way to describe a classical indivisible stochastic (classical) process! Is our conception of reality saved by this discovery? Well, not exactly; indivisible processes are equally as non-intuitive as super-position, entanglement, and all that in quantum theory. But it does seem to hint at an alternative way to explore quantum theory using the basic constructions of classical theory we are so accustomed to.

So let's see; in Quantum theory probabilities are the squares of the complex amplitudes.

i.e if $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ then we will find state $|0\rangle$ with probability α^2 and state $|1\rangle$ with probability β^2
viz,

$$p(|0\rangle) = \alpha^2$$

Now, $\Gamma_{ij}(t)$ are just time dependent probabilities of some stochastic system. Each one is a positive real number between 0 and 1 which means we are more than welcome to express them as

$$\Gamma_{ij}(t) = |\theta_{ij}(t)|^2$$

This is just mathematical fact.

Here we pick $\theta_{ij} \in \mathbb{C}$ but they could be $\mathbb{R}$ or whatever else we may want.

Well just $\theta_{ij}(t)$ are the elements of an $N \times N$ matrix since $\Gamma(t)$ was.

Furthermore $\sum_i \theta_{ij}(t) = 1$

If $\theta(t)$ is a complex matrix then this matrix acts on a vector space spanned by, say, basis vectors

$$\{\vec{e}_i\}$$

e.g.

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

or we can call them $|e_1\rangle, |e_2\rangle$, etc.
In particular, $e_i^\dagger e_j = \delta_{ij}$ or $\langle e_i \mid e_j\rangle = \delta_{ij}$
Additionally, the outer products define "projector Matrices"

$$e_i e_i^\dagger \equiv P_i$$

is a "projector"
e.g

$$e_1 e_1^\dagger = \begin{bmatrix} 1 & 0 & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots \\ \vdots & & \ddots & 0 \end{bmatrix} = |e_1\rangle\langle e_1|$$

Therefore $\theta(t)$ defines a Hilbert space! There's no Hilbert space anywhere from $\Gamma(t)$. The Hilbert space "comes from" writing

$$\Gamma_{ij}(t) = |\theta_{ij}(t)|^2$$

Using these basis vectors $\vec{e}_i = |e_i\rangle$ we have

$$\Gamma_{ij} = \mathrm{tr}(\theta^\dagger(t)|e_i\rangle\langle e_i|\theta(t)|e_j\rangle\langle e_j|)$$

$$\sum_k \langle e_k|\theta^\dagger(t)|e_i\rangle\langle e_i|\theta(t) \mid e_j\rangle\langle e_j \mid e_k\rangle$$

but $\langle e_j \mid e_k\rangle = \delta_{jk}$ so that the sum over k collapses into just the j term and we have

$$\langle e_j|\theta^\dagger(t)|e_i\rangle \langle e_i|\,\theta(t)|e_j\rangle = |\theta_{ij}(t)|^2 = \Gamma_{ij}(t)$$

which is almost exactly how we compute probabilities in quantum theory!

Okay, but at this point this is merely a cute notational trick. We havent said $\Theta(t)$ is unitary and we fundamentally require unitary matrices in quantum theory.

So far we have an indivisible stochastic system defined by $\Gamma_{ij}(t)$ and we took the square root and happened to stumble upon a Hilbert space. But so what? Hilbert spaces occur in all sorts of places; this is not necessarily QM. We'd be really excited if $\Theta(t)$ happened to be "Unitary". So, let's narrow down the class of non-Markovian indivisible stochastic processes further and force the matrices we end up with to be unitary.

Definition: a matrix is called "uni-stochastic" or doubly stochastic if there exists a U such that $\mathscr{U}^\dagger\mathscr{U} = \mathbb{1}$ and

$$\Gamma_{ij} = |\mathscr{U}_{ij}|^2$$

Every matrix like this is "doubly stochastic" meaning each column AND each row sum to 1. Doubly stochastic matrices are very common and it's not such a specific of major limitation to make.

$$\sum_i \Gamma_{ij} = 1 = \sum_j \Gamma_{ij}$$

This means that IF Γ_{ij} is uni-stochastic then we wind up with a unitary Hilbert space!

This then implies we have all the quantum theory we developed so far! Recall, we began with the claim that Quantum objects are
vectors in Hilbert space such that they evolve under unitary transformations (in order to conserve probability - the Born rule).

But we had NO idea where these abstract vectors/spaces came from. We merely said "this is what works". Now, thanks to Barandes, we find that quantum states might/can come from some kind of indivisible uni-stochastic process over some (potentially unknown or hidden) standard configuration space! It implies that all the weirdness of quantum theory is due to the indivisibility of time. It seems materialism might be saved in this view since only distinct configurations exist for the system. Unlike standard quantum theory a system cant be in a superposition of classically observed states/configurations!

In particular our wavefunction $\Psi = \mathscr{U}(t)|e_j\rangle$ for some specific configuration j where $\mathscr{U}$ cones from $\Gamma_{ij}(t) = |\mathscr{U}_{ij}|^2$

But if this is true then a candidate $\Gamma_{ij}(t)$ cannot be divisible. We showed earlier that the things we call quantum are such that the probabilities don't satisfy divisibility. That is to say that there are interference terms which we claimed were due to the level of indivisibility of Γ.

Therefore we postulate the following:

Indivisible Stochastic Systems $\Longleftrightarrow$ Quantum Theory

Let's see if we get our experimentally verified interference starting from an indivisible, uni-stochastic process.

We can define a matrix

$$\mathscr{U}\left(t \leftarrow t'\right) = \mathscr{U}(t)\mathscr{U}^\dagger\left(t'\right)$$

viz, $\mathscr{U}\left(t, t'\right) = e^{iHt}e^{-iH^\dagger t'} = e^{iH(t-t')}$

which is our typical quantum time evolution (remember we can subdivide time in the hilbert space associated with $\mathscr{U}$!)

then, rearranging, we have

$$\mathscr{U}(t) = \mathscr{U}\left(t \leftarrow t'\right)\mathscr{U}\left(t'\right)$$

Now,

$$\Gamma(t) = \overline{\mathscr{U}(t)} \odot \mathscr{U}(t)$$

where $\odot$ is defined by $\Gamma_{ij} = |\mathscr{U}_{ij}|^2$ and is called the "Schur-Hadamard Product". You simply multiply the matrices term by term rather than our normal matrix product.

Then, $\Gamma(t) = \overline{\mathscr{U}(t)} \odot \mathscr{U}(t) = \overline{\mathscr{U}\left(t \leftarrow t'\right)\mathscr{U}\left(t'\right)} \odot \mathscr{U}\left(t \leftarrow t'\right)\mathscr{U}\left(t'\right)$

In terms of the matrix elements this is

$$\begin{aligned}
\Gamma_{ij}(t) &= \sum_{k}^{N} \left|\mathscr{U}_{ik}\left(t \leftarrow t'\right)\right|^2 \left|\mathscr{U}_{kj}(t)\right|^2 \\
&+ \sum_{k\neq l}^{N} \overline{\mathscr{U}_{ik}\left(t \leftarrow t'\right)\mathscr{U}_{kj}\left(t'\right)}\mathscr{U}_{il}\left(t \leftarrow t'\right)\mathscr{U}_{l_j}\left(t'\right) \\
\Gamma_{kj}\left(t'\right) &= \left|\mathscr{U}_{kj}\left(t'\right)\right|^2 \\
\Gamma_{ik}\left(t - t'\right) &= \left|\mathscr{U}_{ik}\left(t \leftarrow t'\right)\right|^2
\end{aligned}$$

But the first term in $\Gamma_{ij}(t)$ is just what we would get if $\Gamma_{ij}(t)$ were divisible!

$$\sum \left|\mathscr{U}_{ik}\left(t \leftarrow t'\right)\right|^2 \left|\mathscr{U}_{k_j}\left(t'\right)\right|^2 = \Gamma\left(t \leftarrow t'\right)\Gamma\left(t'\right)$$

and so,

$$\Gamma(t) - \Gamma(t \leftarrow t')\Gamma(t') = \sum_{k\neq l} \overline{\mathscr{U}_{ik}\left(t \leftarrow t'\right)\mathscr{U}_{k_j}\left(t'\right)}\mathscr{U}_{il}\left(t \leftarrow t'\right)\mathscr{U}_{l_j}\left(t'\right)$$

but $\Psi_k\left(t'\right) = \mathscr{U}_{kj}\left(t'\right)$

Therefore, this term $\Gamma_{ij} - \left[\Gamma\left(t - t'\right)\Gamma\left(t'\right)\right]_{ij}$ is precisely the interference we get from quantum systems!!!

Therefore, it appears that indivisible Uni-stochastic processes are identical to quantum mechanics with its Hilbert space and propagators. We haven't proved this just yet but it seems plausible. We still have to describe measurement as well as composite systems. This was done by Barendes (2023). He finds that we may replace the axioms of quantum theory with the axioms of indivisible uni-stochastic processes.

Indivisible stochastic system are WEIRD. There is no notion of time ordering in this type of non-Markovian process. We aren't allowed to say $A \to C \equiv (A \to B) \cap (B \to C)$ which is our everyday classical experience. (Barendes shows that divisibility emerges from the limit of MANY measurements of an indivisible system.)

Evidently the weirdness of quantum theory is replead by the weirdness of indivisibility. However, this represents a MAJOR step forward in the foundations of quantum theory and the field is wide open for exploratio.