MATH 590-2: Quantum Mechanics and String Theory

Reese Lance

Fall 2022

Abstract

Graduate course in Quantum Mechanics and Strings taught by Paul Aspinwall at Duke. Notes are handwritten during lecture then typeset later. Any comments, concerns, questions, corrections, or communications of any type are encouraged to be directed to my email. This is not a politness request: if you have any feedback about these notes, I really want to hear it. These notes are primarily a documentation of my personal learning journey while following along with the class: There is *a lot* of material in this document that did not come from the lecture, and some of the lecture material may not have been included in these notes. As such, any errors found in this document are assumed to be introduced by me. Nevertheless this should provide some non-zero utility for any and all readers, primarily my future self. This is a math class, but it is meant to be accessible to math undergrads, as well as physics students, so we don't get to treat things as generally or as rigorously as possible at all times. Since rigor and generality are my interest, I will try to provide some generalizations and mathematically rigorous statements for some of the things we are covering, when I am able. These statements will be in the gray-ish boxes, and none of the content in these boxes comes from the course itself, but are sourced mainly from my brain, but also the internet, and the various mathematical physics textbooks I've managed to collect over the years. As such, they are especially susceptible to contain mistakes.

Table of Contents

CHAPTER 1

I. Roadmap and the Classical String

Lecture 1, Aug 29.

The jumping off point of string theory comes in the form of a question. What if the fundamental building blocks of physics is not point particles? The next best guess is instead of a 0-dimensional point, it should be some kind of 1-dimensional object, which could be a string. But how do we describe the motion of such a string? The first guess is to approximate a string as a series of finitely many point particles, which are bound in some sort of potential to each other. But this is not "fundamental". So we may want to start with such a model, then "get rid of the points"[1], and we must do so in a way which is "invariant under reparameterization". This is a difficult constraint to meet, and we will see that it will only work in 26 dimensions, a famous fact.

The usual roadmap to learn string theory for a physicist is to study basic mechanics, in the style of Newton, then classical mechanics via the Lagrangian formulation, then quantum mechanics, quantum field theory, general relativity, and then you are ready to begin your string theory journey[2]. However, this is a very time consuming prospect, as you could be a second or third year PhD student by the time you finish[3] We will follow a somewhat streamlined version of this path, learning only what is absolutely necessary. For example, every undergrad physics students learns about the hydrogen atom in a quantum mechanics class. We will be skipping this completely.

Let's begin with classical mechanics: We have the singular defining equation from Newton, $\vec{F} = m\ddot{\vec{x}}$. Consider the special case when $\vec{F}$ is a conservative vector field[4], i.e. $\vec{F} = -\nabla V$, for $V(x_i, t)$ a scalar field which we think of as representing some sort of "potential energy". Define

$$E := \frac{1}{2}m\dot{\vec{x}}^2 + V$$

[1] I suspect this involves taking some kind of limit as the number of points goes to infinity?

[2] This is exactly what I did in my undergrad, though in a slightly different order, so I can attest this is what people do.

[3] And this is assuming you are a physics student. You will have no time to do all of this if you are a math PhD student.

[4] Some pertinent examples from physics would be gravitational force, Coulomb forces, etc. These forces are characterized by the fact that it doesn't matter what path you take to get to your position, the force exerted on you is the same.

where the square means dotted with itself. Then

$$\frac{dE}{dt} = m\dot{\vec{x}} \cdot \ddot{\vec{x}} + \frac{dV}{dt} = \dot{\vec{x}}(m\ddot{\vec{x}} + \nabla V)$$

But by Newton's equations, the portion in the parenthesis is 0. So for $\vec{F}$ a conservative force field, the "energy", E, is a constant. For a suitable physical situation, we can consider E as the sum of kinetic and potential energy.

In early physics classes, we usually phrase problems in terms of either forces (setting up force diagrams and the like), or in terms of conservation of energy. This can quickly become very complicated though. It is easy to devise some twisted construction of pulleys, weights, and levers which will become too difficult to analyze with these tools.

This is where the Lagrangian formulation of classical mechanics becomes useful.

Setting of Classical Mechanics: To understand this mathematically: We want to work on a Riemannian manifold, M, equipped with the Levi Civita connection, which serves as the congfiguation space for our system. From the manifold point of view, the central objects are paths in the trajectory space $\gamma : [a, b] \to M$ which represent the evolution of the system. With the Riemannian metric, we may consider the gradient as an element of the tangent bundle rather than cotangent, since any inner product induces an isomorphism $V \cong V^*$, and the Levi Civita connection allows us to differentiate "along the path γ". In this case, the equation $F = ma$ takes the form

$$F_\gamma = m\nabla_{\gamma'}\gamma'$$

For concreteness we can take $\mathbb{R}^n$ with the standard metric and trivial connection. From this viewpoint, locally, the generalized coordinates q_i are those coming from the base manifold, and the $\dot{q}_i$ coordinates are those belonging to the tangent vectors. In this way, we view a time-independent Lagrangian as being a function on the tangent bundle, i.e. $L \in C^\infty(TM)$. If one desires time dependence, $L \in C^\infty(TM \times \mathbb{R})$. For the example below, we consider the setting $T\mathbb{R}^3 \times \mathbb{R}$, with coordinates $(x_1, x_2, x_3, \dot{x}_1, \dot{x}_2, \dot{x}_3)$, and $L : T\mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ as described. The technical global definitions in the setting of a general tangent bundle are much more involved. For example, how to globally define a derivative of L so that we may write down Lagrange equations? It can be done, this is known as the fiber-wise derivative. This is a certain transformation $TM \to T^*M$, which we will see later, called the Legendre transform. It is the transformation from Lagrangian to Hamiltonian mechanics. The natural setting for Lagrangian mechanics is TM, and the natural setting for Hamiltonian mechanics is T^*M, leading to the study of symplectic geometry. Mathematically, this explains why the momentum coordinates live in the cotangent bundle, as they are the image of the velocities through the Legendre transform. Details in the next section.

We will use generalized coordinates q_i. And if we are lucky, we can find some special function

$$L(q_i, \dot{q}_i, t)$$

called the Lagrangian, such that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \qquad \text{for all } i$$

These equations are known as the Lagrangian equations.

Example: Let $L = \frac{1}{2}m\dot{\vec{x}}^2 - V$. Note that this differs from the energy by a minus sign, for $i \in \{1, 2, 3\}$. Then we can compute[5]

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = \frac{d}{dt}m\dot{x}_i = m\ddot{x}_i$$

$$\frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i}$$

Then applying the Lagrange equations:

$$m\ddot{x}_i + \frac{\partial V}{\partial x_i} = 0$$

which is exactly the Newton equations.

The calculation above, while suggestive that the Lagrange equations are a good thing to study, is very ad hoc. How did the people who wrote this down initially (probably Lagrange) know it was a good thing to study? To answer this, consider the action

$$S := \int_{t_0}^{t_1} L dt$$

for a system evolving from time $t = 0$ to time $t = 1$. Suppose the action were to vary slightly: That is, suppose the system took a slightly different path through the configuration space, with coordinates q_i, and suppose we forced the endpoints, t_0 and t_1, to stay

[5]For this calculation we are assuming that x and $\dot{x}$ have no relation, and that V depends only on the coordinates x_i and t, not the $\dot{x}_i$. I believe this makes sense. For example, gravitational potential only knows about position, not velocity. It just depends where you stand with respect to the massive object.

fixed. This would correspond to an infinitesimal change in the action, given by[6]

$$\begin{aligned}
\delta S &= \int_{t_0}^{t_1} \delta L dt \\
&= \sum_i \int_{t_0}^{t_1} \left(\left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta \dot{q}_i + \left(\frac{\partial L}{\partial q_i} \right) \delta q_i \right) dt \\
&= \sum_i \cancel{\frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_{t=0}^{t=1}} - \sum_i \int_{t_0}^{t_1} \left(\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} \right) \delta q_i dt \\
\Rightarrow \partial S = 0 &\iff \text{The Lagrangian equations are satisfied}
\end{aligned}$$

where the cancellation occurs because the boundary is fixed, so there is no variation at the boundary, and the third equality (the same line as the cancellation) is obtained via integration by parts, whatever that may mean in this variational setting. To verify the other direction of implication: This $\partial S = 0$ holds for all variations. If the above sum of integrals is 0, then we have something of the form $\int f \cdot g = 0$ for all g, where f is the Lagrange equations and g is the variation δq_i. This implies $f = 0$ (think of the non-degeneracy of the inner product in L^2)[7].

So a trajectory through the configuration space satisfies the Lagrangian equations iff it is an extrema of the action, usually a minimum.

The concept that the system's time evolution should be determined by a stationary state of the action functional is known as Hamilton's principle of least action[8]. It is not guaranteed that we will always find a minimum, but it seems like those are the cases we will be interested in for the most part.

Onto Hamiltonian mechanics. We want a sort of 'coordinate change":

$$p_i := \frac{\partial L}{\partial \dot{q}_i}$$

[6]To make this calculation rigorous, one needs the machinery of variational calculus. But for us, we will just follow our nose and trust things to be well defined and work how we might expect them to. We should expect this concept to be difficult to define because we are trying to define the derivative with respect to a "variation of paths". But the space of paths is infinite dimensional, even for paths in one dimension, so we need to specify what it means to differentiate in this context.

[7]I think this statement is somehow correct but I am still skeptical of my own reasoning. This doesn't hold for all functions g, this holds for all variations δq_i. Maybe if I knew more calculus of variations I could say something definitive.

[8]By the way, I still have no intuition for why this principle should be true. Nature should somehow always choose the configuration which minimizes the action functional? Let's say the action is given by the $K - V$. Nature wants the kinetic energy to be close to the potential energy? Any physicists reading please provide some insight.

Legendre Transformation on Manifolds: There is a more general notion of the Legendre transform between any vector bundle and its dual. In our case, we will specialize to the (co)tangent bundle, but it should be clear how to generalize. The Legendre transform is a smooth map of smooth manifolds $TM \to T^*M$ defined by considering, for any $x \in M$, $L|_{T_xM} : T_xM \to \mathbb{R}$. Then $d(L|_{T_xM}) : T(T_xM) \to T\mathbb{R}$. What exactly is $T(T_xM)$? Think of the simple example of $f : M \to \mathbb{R}$. We can think of df as a map $M \to T^*M$ by sending $p \mapsto \big(v \mapsto df_p(v)\big)$, the directional derivative at p in the direction v. Equivalently we are considering df as a one-form on M, a smooth section of $\Omega^1(M)$. Thus we may consider $d(L|_{T_xM})$ as a map

$$d(L|_{T_xM}) : T_xM \to T^*(T_xM)$$

A one-form is a section of this bundle $T^*(T_xM) \to T_xM$, so the element $d(L|_{T_xM})(v)$ is an element of the fiber over v, $T^*_v(T_xM) \cong T^*_xM$. Since it sends fiber to fiber, this defines a bundle morphism $TM \to T^*M$, denoted as $\mathbf{F}L$, where L is the Lagrangian.

In the case of $L = \frac{1}{2}m\dot{\vec{x}}^2 - V$, we have, if the system obeys the Lagrange equations,

$$p_i = m\dot{x}_i$$

which motivates these new coordinates being referred to as momenta[9].

We define a new function

$$H = \sum_i p_i\dot{q}_i - L$$

the Hamiltonian.

[9]Per our discussion before, we now view these p_i as living in the cotangent bundle.

The Legendre transform reduces to what we just defined: If we choose $M = \mathbb{R}^n$ with coordinates (q_i), or work locally in an arbitrary manifold with local coordinates, and define $L = \frac{1}{2}m\dot{x}^2 - V$, then $\mathbf{F}L : T\mathbb{R}^n = \mathbb{R}^{2n} \to T^*\mathbb{R}^n = \mathbb{R}^{2n}$ and label the coordinates on the domain $\mathbb{R}^{2n}$ as $(q_i, \dot{q}_i)$ and the target as (q_i, p_i). Because $\mathbf{F}L$ is a bundle morphism, it preserves the q_i's, and it sends $\dot{q}_i \mapsto p_i := \frac{\partial L}{\partial \dot{q}_i}$. Note that on $\mathbb{R}^n$, this derivative makes sense. This recovers the general formula given any Lagrangian, and plugging in our particular Lagrangian immediately gives the particular case of $p_i = mq_i$. For this reason, we see the p_i's are coordinates on the cotangent bundle, because they can be interpreted as one-forms.
Under certain nice analytic conditions on L, FL is a diffeomorphism. Define the "Hamiltonian" $H \in C^\infty(T^*M \times \mathbb{R})$ as

$$\begin{aligned} H(q,p) &= p \cdot (\mathbf{F}L)^{-1}(p) - L\big((\mathbf{F}L)^{-1}(p)\big) \\ &= p \cdot \dot{q} - L(\dot{q}) \end{aligned}$$

Then $(\mathbf{F}L)^{-1} = \mathbf{F}H$.
This is a nice way to develop the concept on a manifold, but to actually calculate these things, i.e. to actually compute the Legendre transform of a function and verify the above equation, you need to work with the analytic definiton. Wikipedia is a pretty good source for this, but I'm having some trouble connecting their story on $\mathbb{R}^n$ with what we've developed here. CdA develops it, but over the course of several chapters which I really don't have time to go through right now.

Cannas da Silva's "Lectures on Symplectic Geometry" dive more into the analytic details of this, as well as a more general discussion on arbitrary vector bundles.

Example: Again taking the standard Lagrangian $L = \frac{1}{2}m\dot{\vec{q}}^2 - V(\vec{q})$, we have

$$\begin{aligned} H &= \sum p_i \dot{q}_i - L \\ &= \sum (m\dot{q}_i)\dot{q}_i - \frac{1}{2}m\dot{\vec{q}}^2 + V(\vec{q}) \\ &= \frac{1}{2}m\dot{\vec{q}}^2 + V(\vec{q}) \end{aligned}$$

which we would recognize as $K + V$.

Suppose after plugging in L, we can rearrange H to write it as a function of q_i and p_i[10]. Then

$$dH = \sum_i p_i dq_i + \sum_i dp_i \dot{q}_i - \sum_i \frac{\partial L}{\partial q_i} dq_i - \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i$$

[10]Why can we always do this?

$$= \sum_i \dot{q}_i dp_i - \frac{\partial L}{\partial q_i} dq_i$$

But also,

$$dH \equiv \sum_i \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i$$

$$\Rightarrow \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} \overset{\text{Lagrange Eqns}}{=} -\dot{p}_i, \qquad \frac{\partial H}{\partial p_i} = \dot{q}_i$$

The two equations above are known as Hamilton's equations.

The state of a system is given in terms of p_i and q_i, so a point in phase space has its velocity determined by the Hamilton equations.

Let $f, g(q_i, p_i, t)$. Define the Poisson bracket[11] of f and g by

$$\{f, g\} := \sum_i \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i}$$

TODO: turn the footnote into a section here.

Proposition:

$$\{f, H\} = \frac{df}{dt} - \frac{\partial f}{\partial t}$$

Proof:

$$\frac{df}{dt} \equiv \sum_i \frac{\partial f}{\partial q_i} + \frac{\partial f}{\partial t}$$
$$\overset{\text{Hamilton eqns}}{=} \sum_i \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial t}$$
$$= \{f, H\} + \frac{\partial f}{\partial t}$$

□

In particular, if H has no time dependence, then $\frac{dH}{dt} = \{H, H\} = 0$. The final equality comes by inspecting from the definition[12] of the Poisson bracket that for any f, $\{f, f\} = 0$. We interpret this result as a statement about conservation of energy, in the case that

[11] Every symplectic manifold is canoncially a Poisson manifold: For any f, g, define $\{f, g\} := \omega(X_f, X_g)$, where X_f is the unique vector field such that $\iota_{X_f}\omega = df$. One can check that this indeed defines a Poisson bracket on $C^\infty(M)$. In fact, this bracket turns $C^\infty(M)$ into a Lie algebra. One can work out that in local coordinates, this turns out to give the definition we give subsequently. And this coordinate free definition makes the next lemma much easier to verify.

[12] Or by observing that the symplectic form is a differential form, i.e. antisymmetric.

$H = K + V$.

Let's use the Hamiltonian formulation of mechanics to study the most important system in all of physics, the simple harmonic oscillator, i.e. a system in which $V \sim x^2$. For convenience, we will work in a single dimension, and fix the units so that

$$V = \frac{1}{2}m\omega^2 x^2$$

Then the Hamiltonian is given by

$$H = E = \frac{1}{2}m\dot{\vec{q}}^2 + V(\vec{q}) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

$$E = \frac{1}{2m}\left(p^2 + m^2\omega^2 x^2\right)$$

$$\frac{2E}{m} = \dot{x}^2 + \omega^2 x^2$$

$$\dot{x} = \sqrt{\frac{2E}{m} - \omega^2 x^2}$$

$$\Rightarrow x(t) = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \varphi)$$

where φ is the constant of integration. Alternatively, there exist constants C and $\widetilde{C}$ such that

$$= C\cos(\omega t) + \widetilde{C}\sin(\omega t)$$

where $E = \frac{1}{2}m\omega^2(C^2 + \widetilde{C}^2)$. So this describes the simple harmonic oscillator. It oscillates like a sine function, and the integration constant φ, which would be determined by initial conditions, determines at which phase the oscillation begins.

Now let's move on to our first approximation of a fundamental string, which will be a "violin string": a total of M weights, each of mass ν, a distance d apart, connected by some Hooke's law, and some constant tension when straight, T. If $T = hz$, where z is the length of the string, then

$$V = \frac{1}{2}hz^2 = \int T dz$$

is the energy contained in the spring. When you pluck the string, denote x_k as the vertical displacement of the kth mass, and we assume that the displacement is solely in the vertical direction, which will be the case if you pluck the string very lightly. We assume that x_0 and x_{M+1} are identically 0: The ends of our violin string are fixed, and we fix our coordinate system so that they both are at the same height, and that height is 0. We will also assume that the string's natural/relaxed length is zero[13].

[13] I forget, why do we need to do this? Is this just to emphasize that the thing we really want to think about is the string with infinitely many particles? Anyway this corresponds to having large Hooke's constant, h.

If we imagine the right triangle with points given by the x_i and x_{i-1}, along with the horizontal distance d, then the length of the hypotenuse (which is the length of the string in that portion) is given by

$$\ell_i^2 = \sqrt{d^2 + (x_i - x_{i-1})^2}$$

Thus the total length of the string is

$$\sum_{i=1}^{M+1} \sqrt{d^2 + (x_i - x_{i-1})^2}$$

You can substitute this length into the equation $E = \frac{1}{2}hz^2$, and after some simplification and approximation[14], you will get

$$E = \frac{T}{2d} \sum_{i=1}^{M+1} (x_i - x_{i-1})^2$$

Then we can write the Hamiltonian as

$$\sum_{k=1}^{M} \frac{p_k^2}{2\nu} + \frac{T}{2d} \sum_{k=1}^{M+1} (x_k - x_{k-1})^2$$

$$= \sum_{k=1}^{M} \frac{p_k^2}{2\nu} + \frac{T}{2d}(x_1 x_2 \dots x_M) \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ 0 & -1 & 2 & -1 & & \\ & & & \ddots & & \\ & & & & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix}$$

Denote this scary looking matrix as A. Then, as any mathematician hopes to do when confronted with a matrix, we hope to diagonalize it. Let's compute the characteristic polynomial:

$$c_A(\lambda) = \left| \begin{pmatrix} 2-\lambda & -1 & 0 & & & \\ -1 & 2-\lambda & -1 & & & \\ 0 & -1 & 2-\lambda & -1 & & \\ & & & \ddots & & \\ & & & & 2-\lambda & -1 \\ & & & & -1 & 2-\lambda \end{pmatrix} \right|$$

[14]More details on this simplification can be found in Paul's notes. I like doing math but algebraic symbol pushing is something I'm happy to skip.

Make the coordinate change $2 - \lambda = 2z$. Then the matrix becomes

$$U_M(z) := \begin{pmatrix} 2z & -1 & 0 & & & \\ -1 & 2z & -1 & & & \\ 0 & -1 & 2z & -1 & & \\ & & & \ddots & & \\ & & & & -2z & -1 \\ & & & & -1 & 2z \end{pmatrix}$$

CHAPTER 2

II. The Dynamics of Classical Strings and Continuum Limit

Lecture 2, Aug 31.

We are returning to our discussion of the dynamics of the violin string. Through some analysis, one can show that the eigenvalues of U_M have the form

$$\lambda_n = 2(1-\cos\theta) = 4\sin^2(\theta/2) \equiv 4\sin^2\left(\frac{n\pi}{2(M+1)}\right)$$

for $n = 1, 2, \ldots, M$. Corresponding to the modes of vibration. The eigenvectors can be computed by letting the nth eigenvector be

$$\begin{pmatrix} v_{n,1} \\ \vdots \\ v_{n,m} \end{pmatrix} \in Ker(U_M - \lambda_n I)$$

$$\Rightarrow v_{n,k-1} + v_{n,k+1} = 2\cos\left(\frac{n\pi}{M+1}\right) v_{n,k}$$

Choose $v_{n,1} = A_n \sin\left(\frac{n\pi}{M+1}\right)$, by fixing a normalization. Then $v_{n,k} = A_n \sin\left(\frac{kn\pi}{M+1}\right)$. We want to use Graham-Schmidt to obtain an orthonormal basis, i.e. one which obeys

$$\sum_{k=1}^{M} v_{m,k} v_{n,k} = \delta_{m,n}$$

$$\Rightarrow A_n = \sqrt{\frac{2}{M+1}}$$

Change basis into the eigenbasis:

$$x_k = \sum_n \sqrt{\frac{2}{M+1}} \sin\left(\frac{kn\pi}{M+1}\right) \widetilde{x}_n$$

$$p_k = \sum_n \sqrt{\frac{2}{M+1}} \sin\left(\frac{kn\pi}{M+1}\right) \widetilde{p}_n$$

Now the Hamiltonian can be written nicely:

$$H = \sum_{k=1}^{m} \frac{1}{2\nu}\widetilde{p}_k^2 + \frac{2T}{d} + \sin^2\left(\frac{k\pi}{2M+2}\right)\widetilde{x}_k^2$$

with the Hamilton equations unchanged

$$\frac{\partial H}{\partial \widetilde{q}_n} = -\dot{\widetilde{p}}_n, \qquad \frac{\partial H}{\partial \widetilde{p}_n} = \dot{\widetilde{q}}_n$$

So

$$\widetilde{x}_n = C_n \sin\omega_n t + \widetilde{c}_n \cos\omega_n t$$

where $\omega_n = 2\sqrt{\frac{T}{\nu d}}\sin\left(\frac{n\pi}{2M+2}\right)$ is the frequency over 2π.

Now to approximate the "fundamental string" we need to take the limit $M \to \infty$. In order for the string to not have infinite mass, then we must also let $\nu \to 0$ and we should also let $d \to 0$. If $n << M$, then

$$\omega_n \approx \frac{n\pi}{M+1}\sqrt{\frac{T}{\nu d}}$$

The position on this string is given by kd, $k \in \{0, \ldots, M+1\}$. A sine wave of wavelength λ is thus $\sin\left(\frac{2\pi kd}{\lambda}\right)$.

$$x_k = \sum_n \frac{2}{M+1}\sin\left(\frac{kn\pi}{M+1}\right)\widetilde{x}_n$$

$$\lambda_n = \frac{2(M+1)d}{n}$$

and the velocity of sound is

$$v = \frac{\omega_n}{2\pi}\lambda_n = \sqrt{\frac{T}{\mu}}$$

where $\mu = \nu/d$ is the mass per unit length. Assume we have picked materials such that the velocity of sound is 1 and the length of the string is π, so that $d = \frac{\pi}{M+1}$.

Let $\sigma = kd$ denote the location on a spring. We have $\omega_n = n$, so $x(\sigma, t) = \sum_n A_n e^{-int}\cos(n\sigma)$. If you want x to be real, you need to require that $A_n = A_n^*$[1].

Now we want to consider a closed string. There's a couple problems in doing so. If we clamp it down attached to itself, then the natural length of the springs being 0 implies that this entire loop will contract to 0. And because the two sides are no longer clamped down to something external, the closed loop string can float around in space, so we should keep

[1] I don't think I understand this comment. Doesn't the exponential term make this not real, independent of whatever A does?

track of something like a center of mass motion. If we carry out a similar analysis to the open string, we will find

$$H = \sum_{n=1}^{M} \frac{p_k^2}{2\nu} + \frac{T}{2d}(x_1 \dots x_M) \begin{pmatrix} 2 & -1 & 0 & & & -1 \\ -1 & 2 & -1 & & & \\ 0 & -1 & 2 & -1 & & \\ & & & \ddots & & \\ & & & & -2 & -1 \\ -1 & & & & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix}$$

where now the -1's join the join the ends of the string together. We again define

$$T_M(z) = \frac{1}{2}\det \begin{pmatrix} 2z & -1 & 0 & & & -1 \\ -1 & 2z & -1 & & & \\ 0 & -1 & 2z & -1 & & \\ & & & \ddots & & \\ & & & & -2z & -1 \\ -1 & & & & -1 & 2z \end{pmatrix}$$

It turns out that

$$T_M(z) = \frac{1}{2}\left(U_M(z) - U_{M-2}(z)\right)$$

The T_M's are referred to as the Chebyshev polynomials of the first kind. They also obey the relation

$$T_M(\cos\theta) = \cos(M\theta)$$

The characteristic polynomial of the matrix appearing in the Hamiltonian is given by

$$c_A = 2\left(T_M\left(1 - \frac{\lambda}{2}\right) - 1\right)$$

So the eigenvalues are

$$\theta = \frac{2n\pi}{M}, \qquad \cos\theta = 1 - \frac{\lambda}{2}$$
$$\Rightarrow \lambda = 2\left(1 - \frac{\cos(2n\pi)}{M}\right) = 4\sin^2\frac{n\pi}{M}$$

for $n = 0, \dots, M-1$. Now we have a $\lambda = 0$ eigenvalue. Notice

$$\lambda_{M-n} \equiv 4\sin^2\frac{(M-n)\pi}{M} = 4\sin^2\left(\frac{M\pi}{M} - \frac{n\pi}{M}\right) = 4\sin^2\left(\pi - \frac{n\pi}{M}\right) = 4\sin^2\frac{n\pi}{M} \equiv \lambda_n$$

where we have used the fact that $\sin^2(x-\pi) = (\sin(x-\pi))^2 = (-\sin(x))^2 = \sin^2(x)$. This implies that the eigenspaces will be 2-dim[2]. Noting this, it is convenient to assume

[2]We found two eigenvectors, but you should also show they are linearly independent. We know from LA that eigenvectors corresponding to distinct eigenvalues are linearly independent, but we just showed these eigenvalues are not distinct. This comes because sin and cos are orthogonal.

M is odd, and let the eigenvalues be indexed by $|n| < M/2$, so that the eigenvalues for $\pm n$ coincide. When $n \neq 0$, we have the two normalized eigenvectors

$$\sqrt{\frac{2}{M}} \sin\left(\frac{2nk\pi}{M}\right), \qquad \sqrt{\frac{2}{M}} \cos\left(\frac{2nk\pi}{M}\right)$$

with wavelength given by Md/n, where d is the distance between two neighboring molecules. When $n = 0$, this is interpreted as the string moving with no potential, so it moves as a free particle with constant velocity along its center of mass.

The general solution can be written

$$x_k = \vec{x} + \vec{v}t + \sum_{|n|<M/2, n\neq 0} C_n \exp\left(i\omega_n t + \frac{2nk\pi i}{M}\right) + \widetilde{C}_n \exp\left(i\omega_n t - \frac{2nk\pi i}{M}\right)$$

where $\vec{x}$ is the center of mass at $t = 0$ and $\vec{v}$ is the velocity of the COM. Again taking the continuum limit and letting the length of the string be 2π, we get

$$x(\sigma, t) = \vec{x} + \vec{v}t + \sum_{n\neq 0} \left(A_n e^{-in(t+\sigma)} + B_n e^{-in(t-\sigma)}\right)$$

The difference between the closed loop string and the violin string are the presence of the $n = 0$ mode and the existence of left and right movers[3]. The A_n term moves left and the B_n term moves to the right. There is also an open string, which is not clamped at all, and the DN (Dirichlet-Neumann) string, which is clamped at one end. We will not study these as closely, so we don't engage with them here, but Paul's notes have a detailed account of their dynamics.

[3]I'm a bit confused here. Now that I think about it, shouldn't there also be left and right movers in the violin string, since a right mover would reflect off the other end and become a left mover. Maybe the idea is it would destructively interfere with the right mover?

CHAPTER 3

III. Intro to QM

Lecture 3, Sept 5.

To start off with Quantum Mechanics, we need to assume some physically motivated "postulates". A mathematician may refer to these as axioms.

Definition: A Hilbert Space is a vector space with an inner product, $\langle \cdot, \cdot \rangle$, which is complete[1] with respect to that inner product.

Typically our Hilbert spaces are infinite dimensional, but that is not necessary. Further, we take our Hilbert spaces to be complex inner product spaces, meaning the underlying VS is complex, and the inner product is conjugate symmetric rather than symmetric, and it is not bilinear anymore, but linear in only the second term.

Details about Hilbert Spaces: Note that if a Hilbert space $(\mathscr{H}, \langle \cdot, \cdot \rangle)$ is finite dimensional, then the completeness condition is automatic: Choose an orthonormal basis of V, $\{v_i\}$, with respect to $\langle \cdot, \cdot \rangle$, which we may do via Graham-Schmidt, and because the inner product is always non-degenerate. Then define $T : V \to \mathbb{C}^n$, with $T(v_i) = e_i$. Denoting the standard inner product on $\mathbb{C}^n$ as $\langle u, v \rangle_{\mathbb{C}^n}$, we have

$$\langle T(v_i), T(v_j) \rangle_{\mathbb{C}^n} \equiv \langle e_i, e_j \rangle_{\mathbb{C}^n} = \delta_{ij} = \langle v_i, v_j \rangle$$

So T is an isometry, and thus preserves completness, so V is complete.

Another intuitive way (which I believe can be made rigorous) to see this result is that every f.d. VS is isomorphic to $\mathbb{C}^n$, with a possibly different inner product. But all norms on f.d. VS are equivalent, so both norms must be complete, since the standard norm is.

[1] All Cauchy sequences in the space converge to a point in the space. In other words, there are no "holes" in the space.

The key example we will use in quantum mechanics is L^2, the space of square integrable functions:

Definition: The L^p space of a measurable space (X, Σ, μ) is defined as the space of μ-measurable functions such that

$$\left(\int_X |f|^p d\mu\right)^{\frac{1}{p}} < \infty$$

One can check that this is a vector space, and the definition $\|f\|_p := \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}$ turns $L^p(X)$ into a normed space. Due to the Born rule, we will want to consider $p = 2$ usually. This is the only p so that the norm actually comes from an inner product: for $f, g \in L^2(X)$,

$$\langle f, g\rangle := \int_X f(x)\overline{g(x)}d\mu(x)$$

But the details of that entire sentence are quite involved to verify.
In QM we also require our Hilbert spaces to be separable, which means they admit a countable orthonormal basis, as opposed to uncountable. For example, $L^2(\mathbb{R})$ is separable, with countable basis given by e^{inx}. But it turns out most spaces we could consider are separable anyway, so often this is not a concern.

Now we introduce the Dirac notation: A vector $v \in \mathscr{H}$ is denoted as $|v\rangle$, and an element of the dual $w \in \mathscr{H}^*$ is denoted $\langle w|$. We drop the parenthesis so that $\langle w|v\rangle \in \mathbb{C}$.

Given $v \in \mathscr{H}$, we may consider its image through the dual map $*_{\langle\cdot,\cdot\rangle} : V \to V^*$, $u \mapsto \langle u, \cdot\rangle$. We then denote $\langle u|v\rangle \equiv \langle u, v\rangle$.

TODO: natural isomorphism $** \cong Id$ on the category of vector spaces (maybe)

Note that $a|u\rangle \mapsto \bar{a}\langle u|$, since we have a complex inner product space. So this map is not linear but anti-linear. For real vector spaces it is linear.

Given a basis of $\mathscr{H}$, we can think of a vector/ket as a column vector and a covector/bra as a row vector with conjugated components.

Postulate 1 (State Space): Given a physical system with some Hilbert space $\mathscr{H}$, a "state" of the system is given by a ray in the Hilbert space:

Definition: A ray[2] in $\mathscr{H}$ is an equivalence class of non-zero vectors under the equivalence relation

$$|v\rangle \sim |w\rangle \iff |v\rangle = \lambda|w\rangle, \qquad \lambda \in \mathbb{C}^*$$

In other words, a state is an element of $\mathbb{C}P^{\dim \mathscr{H}}$. We will see why this is the appropriate definition shortly.

Definition: Let A be an operator on $\mathscr{H}$. $A^\dagger$ is the adjoint to A if

$$\langle A^\dagger u, v\rangle = \langle u, Av\rangle \quad \forall\, u, v$$

The dagger induces a right action of $\mathscr{H}$ on $\mathscr{H}^*$:

$$\langle u|\, A := \langle A^\dagger u|$$

So that

$$(\langle u|\, A)\, |v\rangle = \langle A^\dagger u|\, |v\rangle \equiv \langle A^\dagger u, v\rangle = u, Av\rangle \equiv \langle u|\, (A\, |v\rangle)$$

thus writing

$$\langle u|\, A\, |v\rangle$$

is unambiguous. In terms of a basis, $A^\dagger$ is the conjugate transpose of A, as we showed on the homework. Further,

$$\langle u|\, A\, |v\rangle = (\bar{u}_1\ \bar{u}_2 \dots) A \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}$$

Definition: A linear operator A on $\mathscr{H}$ is called self-adjoint if $A = A^\dagger$. It is called unitary if $A^\dagger = A^{-1}$.

Theorem[3] **(Spectral Theorem):** *If A is a self-adjoint operator on a finite dimensional vector space, then there is an orthonormal basis of eigenvectors and all eigenvalues are real.*

Postulate 2 (Observables): An observable, as in something you would measure in a lab, corresponds to a self adjoint operator on the Hilbert space.

[2] Be careful about using your intuition for what this would look like. If the field is complex numbers, a ray in this sense of the word will not just be a straight line, since multiplication by a complex number can rotate the plane.

[3] State the more general version of spectral theorem for normal matrices? Also should discuss inf dim since that's what we need.

Some Functional Analysis: Definition: Given a bounded linear operator T on a (possibly infinite dimensional) vector space, the spectrum of T is the set of all $\lambda \in k$ such that the linear operator $T - \lambda I$ is not a bijection, denoted $\sigma(T)$. Note that if V is finite dimensional, $\sigma(T)$ corresponds exactly to the set of eigenvalues.

In the infinite dimensional case, it is not guaranteed that any linear operator will have an eigenvalue, even when the field is algebraically closed. This is because the fundamental theorem of algebra holds only for finite degree polynomials.

As an example, consider a countably infinite dimensional vector space, with basis $\{e_i\}$ and consider the operator A sending

$$e_i \mapsto e_{2i}$$

Clearly this map does not have any eigenvalues. However, the map $(A - I)$ sends

$$\sum_{i=1}^{\infty} a_i e_i \mapsto \sum_{i=1}^{\infty} a_i e_{2i} - \sum_{i=1}^{\infty} a_i e_i$$

which is not surjective. For example, it doesn't hit e_1: To get e_1, you must have $a_1 = -1$. So you get sent to $e_1 - e_2$. You need to kill the e_2 term, which you can do by letting $a_2 = -1$. Then you get mapped to $e_1 - e_2 - (e_4 - e_2) = e_1 - e_4$. To kill e_4, you let $a_4 = -1$, which is mapped to $e_1 - e_8$, and so on. You must have infinitely many a_i's nonzero to map onto e_1, which is not allowed. So this map is not surjective, and $1 \in \sigma(A)$. In general, if you have a linear operator which you know has no eigenvalues, then the only way for λ to be in the spectrum is if $A - \lambda I$ fails to be surjective. It cannot fail to be injective.

Though A may not have eigenvalues, there is a nice result which generalizes the concept for infinite dimensional spaces over $\mathbb{C}$:

Proposition: *If A is a bounded (with respect to the norm on $\mathbb{C}$) linear operator on an infinite dimensional vector space over $\mathbb{C}$, then $\sigma(A)$ is always non-empty.*

Proof: Define the resolvent of A, $\rho(A) := \sigma(A)^C \subset k$. This is the set of λ for which $A - \lambda I$ is invertible. We also define the resolvent function

$$R : \rho(A) \to End(V)$$

$$\lambda \mapsto (A - \lambda I)^{-1}$$

If $\sigma(A)$ is empty, then R is defined on all of $\mathbb{C}$ and is bounded. It can also be shown to be holomorphic and thus entire (I'm not even sure I know what this means in this setting), so by a generalized Liouville theorem, R is constant. But $R(\infty) = 0$ (non-trivial to show), so $R \equiv 0$, which is a contradiction.

□

Postulate 3: The allowed values of the observable correspond to elements of the spectrum of its corresponding self-adjoint operator.

Some More Functional Analysis: This discussion actually matters, it is not just generalization for the sake of it. In QM classes it is often stated that the allowed values of an observable are the eigenvalues of the corresponding self-adjoint operator. But for example, the position operator on $L^2(\mathbb{R})$ sending $f(x) \mapsto xf(x)$ does not have any eigenvalues. There are no functions such that $xf(x) = f(x)$ except for 0. One would be tempted to say that the Dirac delta function satisfies this property, but such a thing does not live in $L^2(\mathbb{R})$, as it is not a function at all, but a distribution. Clearly this cannot be the case because that implies that you can not measure anything for position. Instead we must consider its spectrum, and what is that? We hope to get $\mathbb{R}$ in the end. Note that the discussion and definitions above were only for bounded operators, but it is intuitively clear that position and momentum operators are not bounded, since their range is infinite. So we need to expand our definitions from even what we considered above. But the above was included to guide where we are headed.

Technically an unbounded operator on a Hilbert space $\mathscr{H}$ is just an operator $L : Dom(L) \to \mathscr{H}$, where $Dom(L)$ is the domain of L, a subspace of $\mathscr{H}$. There is nothing in this definition referencing boundedness. One should think of "unbounded" to mean "not necessarily bounded". We will frequently consider cases where $Dom(L)$ is a dense subspace of $\mathscr{H}$, called densely defined. This is an equivalent condition to the existence of an adjoint operator, so we now see why that should be important in QM. In particular, operators which are defined on all of $L^2(\mathbb{R})$ are densely defined. I've read some conflicting things on the internet about whether the position operator is defined over all of $L^2(\mathbb{R})$ or not. For instance this stack exchange post claims it is only defined on the subspace such that $xf(x) \in L^2(\mathbb{R})$, while this pdf claims it is defined on all of $L^2(\mathbb{R})$. I thought it should be the former case, since multiplying by x could take you out of $L^2(\mathbb{R})$, but I don't know anymore. Perhaps these definitions are subtly different.

TODO: Finish the spectrum of an unbounded operator and calculate what the spectrum of position and momentum are.

This postulate tells us that we mainly will be interested in infinite dimensional Hilbert spaces: If we have an observable represented by a linear operator on a finite dimensional VS, then its spectrum is a finite set, since each distinct eigenvalue eats up at least one dimension of the VS. But values that we would want to measure in real life are often represented by real numbers[4], such as position and momentum.

[4]Depending on your philosophical feelings towards spacetime

Suppose we have an orthonormal basis of $\mathscr{H}$[5] and let A represent an observable which is diagonalized in this basis. Any state can be expanded uniquely as $|\psi\rangle = \sum c_i |a_i\rangle$.

Definition: We say a vector $\psi \in \mathscr{H}$ is normalized if $\langle\psi|\,|\psi\rangle = 1$, i.e. $\sum_i |c_i|^2 = 1$. Note that it doesn't make sense to ask if a state of a system is normalized, but you may always pick a normalized representative of the state by choosing any representative and dividing it by its magnitude. By the definition of the equivalence relation, this will be an element of the same equivalence class. However this choice of normalized representative is not unique, as multiplying it by any other phase, $e^{i\theta}$, preserves the norm. In fact any two normalized representatives of the same state must be related by a multiplicative factor of $e^{i\theta}$.

Example (Qubit): A qubit is an element of $\mathscr{H} = \mathbb{C}^2$. Any observable of this system has only two possible values, since the Hilbert space is 2 dimensional, thus there can be at most 2 distinct eigenvalues of any linear operator. In their respective basis, these are 1 and 0, and $|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Postulate 4 (Born rule): If a state is represented by a normalized vector $|\psi\rangle$, and we have an operator A corresponding to an observable, with orthonormal eigenbasis $|a_i\rangle$, then the probability of "measuring A" and getting the value λ_i is

$$|\langle a_i|\psi\rangle|^2$$

The elements of the eigenbasis are also normalized, and thus represent physical states of the system. These are known as "stationary states", states where you know the value of A measured on these states with 100% certainty: you will get λ_i. If your state is a non-trivial linear combination of eigenstates, then the outcome is uncertain. For the qubit:

$$|\psi\rangle = |0\rangle \Rightarrow \text{ You will measure 0}$$
$$|\psi\rangle = |1\rangle \Rightarrow \text{ You will measure 1}$$
$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \Rightarrow \text{ You will measure either 1 or 0 with 50\% probability}$$
$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle) \Rightarrow \text{ You will measure either 1 or 0 with 50\% probability}$$

These are all trivial to check. The general pattern is clear as well: $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$, then the probability of measuring 0 is $|c_0|^2$ and similarly for 1. This was obvious from the moment we stated the Born rule, though, for those used to linear algebra.

If $\mathscr{H} = \mathbb{C}^2 = \mathbb{R}^4$, then you can quotient out by the equivalence relation of rays, i.e. consider the space of rays in $\mathbb{C}^2$ rather than all of the individual points. We killed scaling

[5] this can be done in infinite dimensions as well via Graham-Schmidt, but I don't know the details of how.

by complex numbers, so we deleted a complex dimension. The resulting space has complex dimension 1, and real dimension 2. It turns out to be S^2, the unit sphere. Then you can track the basis vectors through this transformation and draw a nice picture known as the Bloch sphere. This representation is used very frequently in quantum computation.

What is the space of rays? More generally, if we have a Hilbert space of dimension n, then the space of rays in $\mathscr{H}$ is a well known mathematical construction known as $\mathbb{C}P^n$. This is a complex manifold of complex dimension $n-1$, constructed more or less how we described before. Start with $\mathscr{H}$. Take away the point $\vec{0}$. Then quotient by the equivalence relation $v \sim w \iff v = \lambda w$ for some $\lambda \in \mathbb{C}^*$. To see the manifold structure, $\mathbb{C}P^n$ is covered by sets $U_i = \{\vec{z} \in \mathbb{C}P^n \mid z_i \neq 0\}$, viewing $\vec{z}$ as an element of $\mathscr{H}$. This covers because $\vec{0} \notin \mathbb{C}P^n$, so every element must have at least one non-zero entry. But each U_i is just another copy of $\mathbb{C}^n$: define

$$\varphi_i : U_i \to \mathbb{C}P^n$$

$$(z_1, \ldots, z_{n+1}) \mapsto \frac{1}{z_i}(z_0, \ldots, \hat{z_j}, \ldots, z_{n+1})$$

where the hat denotes that z_j has been omitted. As an exercise, show that this map is a homeomorphism: write down the inverse (and make sure it is well defined) and observe that they are both continuous. One must also compute the transition charts and observe that they are smooth, as maps between $\mathbb{R}^{2n}$ and $\mathbb{R}^{2n}$. There is a lot of things to check, but all of them are routine. An even more general construction one could investigate is the Grassmannian, $Gr(n,k)$. Then $\mathbb{C}P^n = Gr(n,1)$, and there are similar manifold constructions for $Gr(n,k)$, but I don't believe the Grassmannian has any use in quantum mechanics specifically, but it does appear frequently in other realms of mathematical physics, particularly the Affine Grassmannian. I do not understand this object yet, though. I have not put in any effort on this front because I am very busy.

Canonical quantization is the assumption of all of the above postulates along with the prescription $[p,q] = i\hbar$, where p and q are momentum and position operators, respectively. The bracket indicates the commutator, $[A,B] := AB - BA$. Note that this prescription only makes sense in an infinite dimensional Hilbert space. If you were to try to define it for a finite dim Hilbert space, you could take trace of both sides, yielding $ni\hbar = 0$, which is false. How could you motivate such a prescription? Notice that if A and B are self-adjoint (as will be the case for us), this implies

$$[A,B]^\dagger = -[A,B]$$

So that if $[A,B]$ is a constant, c, then c must contain a factor of i. This explains the inclusion of i, and $\hbar$ is a constant which corrects units. In natural units, we usually set $\hbar = 1$, among other things.

Note that

$$\begin{aligned}[A, BC] &= ABC - BCA \\ &= ABC - BAC + BAC - BCA \\ &= [A, B]C + B[A, C]\end{aligned}$$

which is a kind of Leibniz rule. This implies

$$[q, p^n] = np^{n-1}[q, p]$$

and more generally, for any f analytic,

$$[q, f(p)] = \frac{\partial f}{\partial p}[q, p]$$

and

$$[g(q), p] = \frac{\partial g}{\partial q}[q, p]$$

It is also "sort of"[6] true that

$$[A, B] = \{A, B\}[q, p] = \{A, B\}i\hbar$$

This the Dirac formulation of canonical quantization.

TODO: **Groenewold's Theorem**, which basically says that quantization is a scam

Classically, we saw

$$\frac{dA}{dt} = \{A, H\} + \partial_t A$$

time translation was generated by bracketing with the Hamiltonian. Thus in QM,

$$i\hbar\frac{dA}{dt} = [A, H] + \partial_t A$$

So we think of operators as evolving in time. This is the Heisenberg picture, so the observables change and the states stay the same. The opposite POV should give the same physics. This leads to the Schrodinger equation and Schrodinger POV: Everything should be determined by "matrix elements"

$$\langle u| A |v\rangle$$

[6] I'd like to find out exactly what this means.

In the Schrodinger picture, the states $|v\rangle$ evolve in time. Let's suppose we want the inner product to remain constant[7]:

$$\begin{aligned}
\frac{d}{dt}\langle u|\, v\rangle &= 0 \\
&= \langle \dot{u}|\, v\rangle + \langle u\,|\dot{v}\rangle \\
\Rightarrow \left(\frac{d}{dt}\right)^{\dagger} &= -\frac{d}{dt} \\
\Rightarrow \frac{d}{dt}\langle u|\, A\,|v\rangle &= \langle u|\left[A, \frac{d}{dt}\right]|v\rangle \\
\Rightarrow i\hbar\,\langle u|\, A\,|v\rangle &= \langle u|\left[A, i\hbar\frac{d}{dt}\right]|v\rangle
\end{aligned}$$

From the Heisenberg picture, we derived

$$i\hbar\frac{d}{dt}\langle u|\, A\,|v\rangle = \langle u|\, i\hbar\frac{d}{dt}A\,|v\rangle = \langle u|\,[A, H]\,|v\rangle$$

Comparing this with the above line, we find the Schrodinger equation:

$$H = i\hbar\frac{d}{dt}$$

[7]Why should we want that? If the states can evolve in time, doesn't it make sense for the probability of transmission from one state to another to also change in time?

CHAPTER 4

IV. Missed lecture.

Lecture 4, Sept 7.

CHAPTER 5

V. Quantization of the Classical String

Lecture 5, Sept 12.

My own brief summary of the lecture I missed:
Last time we had the Hamiltonian for a simple harmonic oscillator, which by definition is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

By the clever introduction of operators $u, u^\dagger$, we could rewrite H as

$$H = \left(u^\dagger u + \frac{1}{2}\right)\hbar\omega$$

If $|\psi\rangle$ is a normalized eigenstate of H, then we found that u and $u^\dagger$ are raising and lowering operators, the sense that they bump up or down eigenvalues of H. There is a ground state, a state which is annihilated by the lowering operator, and a highest energy state, one which is annihilated by the raising operator. In this way, the energy levels of a SHO are quantized. This can be generalized to our violin string situation, which we approximate as having M SHO's. Then the Hamiltonian is of course

$$H = \sum_{n=1}^{M}\left(u_n^\dagger u_n + \frac{1}{2}\right)\omega_n$$

After some analysis, this leads to solutions[1] of the form

$$\hat{x}_n = \sum_{n=1}\frac{1}{\sqrt{n\pi}}\left(u_n e^{-int} + u_n^\dagger e^{int}\right)\sin(n\sigma)$$

it is convenient to normalize the raising and lowering opeartors as

$$\alpha_n = \sqrt{n}u_n, \quad \alpha_{-n} = \sqrt{n}u_n^\dagger$$

for $n \geq 1$. Then the commutation relations are

$$\boxed{[\alpha_m, \alpha_n] = m\delta_{m+n}}$$

[1] TODO: solutions of what?

We will use this a lot. Then the solutions can be written more succintly in terms of the α's, and letting $M \to \infty$,

$$\hat{x}(\sigma, t) = \frac{1}{\sqrt{\pi T}} \sum_{n \neq 0} \frac{1}{n} \alpha_n e^{-int} \sin(n\sigma)$$

We had to leave 0 out because it is not clear what α_0 should mean yet. We can then calculate directly the brackets at the same time

$$[\hat{x}(\sigma, t), \hat{x}(\sigma', t)] = 0$$

We could also consider the same time but with a derivative in one place:

$$[\dot{\hat{x}}(\sigma, t), \hat{x}(\sigma', t)]$$

Just expanding the definitions, one gets

$$= \frac{1}{iT}(\delta(\sigma - \sigma') - \delta(\sigma + \sigma'))$$

with δ taken mod 2π. Then

$$\delta(\sigma + \sigma') = 1 \iff \sigma = \sigma' = 0 \text{ or } \sigma = \sigma' = 2\pi$$

i.e. it only contributes at the ends of the string. At the ends of the strings, $\delta(\sigma - \sigma') = 0$, so only one delta will ever appear. Define $\hat{p} = T\dot{\hat{x}}$. Then away from the ends of the string,

$$[\hat{x}(\sigma, t), \hat{p}(\sigma', t)] = i\delta(\sigma - \sigma')$$

We could also consider the commutator at different times[2]. So the story of the quantized is mostly done. But what about our ground state? We want a state with $\alpha_n |0\rangle = 0$. This would imply

$$\begin{aligned} H|0\rangle &= \sum_{n>0} \left(\alpha_{-n}\alpha_n + \frac{n}{2}\right)|0\rangle \\ &= \sum_{n>0} \left(\frac{\alpha_{-n}\alpha_n}{n} + \frac{1}{2}\right) n|0\rangle \\ &= \sum_{n>0} \frac{1}{n}\alpha_{-n}\alpha_n|0\rangle + \frac{1}{2}\sum_{n>0} n|0\rangle \\ &= 0 + \frac{1}{2}\sum_{n>0} n|0\rangle = \infty|0\rangle \end{aligned}$$

So our ground state has infinite energy! That's not good. Physics is invariant under global shift, so rescale H so that

$$H|0\rangle = 0$$

[2]Spin-statistics theorem is relevant here.

by setting[3]

$$H = \sum_{n>0} \alpha_{-n}\alpha_n$$

and this fixes our problem. However, we have changed the energy of the vacuum, which really matters for certain subjects in physics, particularly GR. So we are actually changing physics. This is one of the ways in which ST appears to be incompatible with GR. So we now have a vacuum state with 0 energy, denoted $|0\rangle$. We also have $\alpha_{-1}|0\rangle$ is the first excited state, with energy $E = 1$ (recall the energy is its H-eigenvalue). We could apply α_{-1} again, or apply α_{-2} to the vacuum. These are not the same thing, but they do have the same energy of 2.

We can do a similar analysis for the closed string. Let the COM coordinate be $\bar{x}$ with momentum $\vec{p}$. The

$$x(\sigma, t) = \bar{x} + \frac{pt}{2\pi T} + \frac{1}{2\sqrt{\pi T}} \sum_{n\neq 0} \frac{1}{n}\left(\alpha_n e^{-in(t+\sigma)} + \widetilde{\alpha}_n e^{-in(t-\sigma)}\right)$$

with $[\bar{x}, p] = i$, $[\alpha_m, \alpha_n] = m\delta_{m+n}$, $[\widetilde{\alpha}_n, \widetilde{\alpha}_m] = m\delta_{m+n}$, so the α's represent quantized left-movers and $\widetilde{\alpha}$'s represent quantized right-movers.

We can also consider displacements in multiple dimensions. But this just corresponds to replacing the displacement x with a vector displacement x^μ. Then

$$H = \sum_{\mu=1}^{D} \frac{(p^\mu)^2}{2\nu} + \frac{1}{2}\nu\omega^2(x^\mu)^2$$

with

$$[\alpha_m^\mu, \alpha_n^\nu] = m\delta_{\mu\nu}\delta_{m+n}$$

So ignoring the problem of the vacuum renormalization, this completes the story of the quantization of the classical string. Now we want to incorporate relativity. But this just corresponds to increasing the dimension of our space and calling the new dimension "time". Remember one key postulate of relativity is that space and time should be treated as equal, so we treat time as just adding another dimension, which we just discussed how to do. We let $x^0 =$ "time" and x^μ be spatial coordinates for $\mu = 1, \ldots, D-1$.

If we want to take this prescription of treating the time dimension as the same as space, then just as our string is allowed to oscillate back and forth in space, it must also be allowed to oscillate back and forth in time. We want to eliminate t from the Hamiltonian.

Note every string defines a world sheet surface, Σ in spacetime by tracking its time evolution. As a 2-manifold, equip Σ with local coordinates τ, σ, and let x^μ serve as the local

[3] I think what is happening is we are rescaling each individual SHO by a finite amount, specifically shifting the potential energy down by $\frac{1}{2}\hbar\omega$, which is perfectly allowed, and if we do it before we take the limit, this only needs to be done finitely many times, which should be kosher.

chart[4] from $\mathbb{R}^2_{\sigma,\tau} \to U \subset \Sigma$.

In our definition of H, we used an inner product to write, for example, p^2. but this has the wrong signature. To be compatible with SR, we know we must use the Minkowski metric,

$$x \cdot y = -x^0 y^0 + \sum x^i y^i$$

Now we introduce the Einstein notation, which is the convention that repeated up and down indices are implicitly summed over, omitting the summation notation. So define the metric $\eta_{\mu\nu} = diag(-1, 1, 1, \dots)$. Then we can raise and lower indices:

$$x_\mu = \eta_{\mu\nu} x^\nu$$

Note

$$\eta^2_{\mu\nu} = 1 \Rightarrow x^\mu = \eta^{\mu\nu} x_\nu$$

where $\eta^{\mu\nu}$ is the matrix inverse of $\eta_{\mu\nu}$, i.e.

$$\eta_{\mu\nu}\eta^{\gamma\beta} = \delta_{\mu,\gamma}\delta_{\nu,\beta}$$

Ultimately, this means

$$[\alpha^\mu_m, \alpha^\nu_n] = m\delta_{m+n}\eta^{\mu\nu}$$
$$\Rightarrow [\alpha^0_m, \alpha^0_n] = -m\delta_{m+n}$$

Also $\|\alpha^0_{-1}|0\rangle\|^2 = -1$? this inner product is not positive definite. It is maybe not clear exactly how, but this problem arises because we allowed the string to vibrate forwards and backwards in time. To fix this, we need to address our time definitions. We could assert $x^0 = \tau$. This is bad, though, because now we are distinguishing time from space. So maybe we could define $x^0 = f(\tau)$. But then we need to ignore reparameterizations. But then we must also ignore spatial reparameterizations, since we treat space and time equally. This means there is now no meaning to "position" on the string. This is what we mean by "fundamental string". There should be no way to distinguish ones position on the string. So to incorporate SR into our quantized string, we need to account for its symmetries.

This begins our discussion of Lie groups, Lie algebras and representation theory. This is a bit more math-y than before.

Definition: A group is a one-object groupoid. I don't want to write out the full definition.

There are levels to how bad a group can be: i) finite is not so bad. ii) countably[5] infinite groups. iii) uncountably infinite groups (but still finite dimensional, such as $\mathbb{R}$. iv)

[4] I'm not sure I understood this part correctly. In class Paul said that this map x^μ need not have any restraints. So it must not be a local chart then, which is required to be a homeomorphism?

[5] As everyone knows, the Monster group is much more tame than $\mathbb{Z}$.

the worst, infinite dimensional groups.

String theory symmetry groups will be of type iv). Consider the Lie group S^1.

Basics of Lie theory: A Lie group is a group object in the category of smooth manifolds. Maybe more usefully:

Definition: A Lie group is a group G, which is also a smooth manifold, such that the maps $i : G \to G$ sending $g \mapsto g^{-1}$ and $\mu : G \times G \to G$ sending $(g_1, g_2) \mapsto g_1 g_2$, are smooth maps, in the sense of smooth manifold theory.

Lie groups are often considered as "symmetry groups" in physics.

Example: S^1 can be considered as a group with elements $\theta \in [0, 2\pi]$ with ends identified. The group law is given by addition of angles mod 2π. S^1 also has a smooth 1-dimensional manifold structure. We will not go through the details here, but the gist is that around any point, a neighborhood just "looks like" $\mathbb{R}$. This can all be made precise without much difficulty. There's many things to check but they are all routine. In fact, S^n for every $n \geq 0$ is a smooth manifold, through a very straightforward generalization of the techniques used in $n = 1$.

Fun fact: Though every sphere is a smooth manifold, not every sphere is a Lie group. One way to see this is to observe that every Lie group is parallelizable, i.e. has trivial tangent bundle. Intuitively this is because each tangent space can be canonically identified with the tangent space at the identity by the push forward of left multiplication, L_{g*}, which is a diffeomorphism because L_g is a diffeomorphism, by definition of a Lie group. So in order for S^n to be a Lie group, we must have TS^n is a trivial vector bundle. Equivalently, it requires admitting n linearly independent vector fields on S^n. In a talk I gave at UNC last year (2021), I proved constructively that this can be done for $n = 0, 1, 3$, and 7 by constructing Clifford algebra representations corresponding to linearly independent vector fields on S^n. In reality, this result, while highly suggestive, and a very fun calculation, doesn't prove anything. Having parallelizable tangent bundle does not, as we will show in a second, imply that the manifold is a Lie group, and my calculation does not show that this can't be done on any other spheres (although I cited a theorem which does say that the strategy I presented is optimal, but the techniques to prove the theorem are more sophisticated.). In fact, it turns out even though $n = 7$ satisfies this criteria, S^7 is not a Lie group, but the other 3 are, and they are the only such. Considering that talk employs only linear algebra and representation theory, that is a pretty good result in my eyes. Notes for this talk can be found in the "Notes" section of my website, which I'm presuming is where you found this document as well.

This group can act on $\mathbb{R}^2$ as rotations about the origin.This is formalized with the language of a group representation, which we will see later. In this case, we have

$$S^1 \to End(\mathbb{R}^2)$$

$$\theta \mapsto R_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

We could take a Taylor series expansion

$$R_\theta \approx 1 + \epsilon X + O(\epsilon^2)$$

where $X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Definition: Given a Lie group G, define the Lie algebra *of G* as the tangent space of G at the identity, T_eG, usually denoted $\mathfrak{g}$.

The Lie algebra of a Lie group: We defined the Lie algebra as T_eG, which by elementary differential geometry is a vector space.

Definition: A vector field on a Lie group, $X \in \mathfrak{X}(G)$, is called left-invariant if, for all $g \in G$,

$$dL_g(X) = X \circ L_g$$

where

$$L_g : G \to G$$
$$h \mapsto gh$$

The above definition is an equality of vector fields. (sanity check for the reader: make sure you know how to plug in a point to both sides of that equation and get out a vector in the tangent space over that point.)

One can check that the above is a vector space. Denote it as $\mathfrak{X}(G)^{L_g}$.

Theorem: There is an isomorphism of vector spaces $T_eG \cong \mathfrak{X}(G)^{L_g}$.

Proof: Define a map

$$T_eG \to \mathfrak{X}(G)^{L_g}$$
$$v \mapsto X^v$$

where $(X^v)_g \equiv (dL_g)_e(v) \in T_gG$. We defined a vector field, but we need to check it is left-invariant:

$$\begin{aligned} & dL_h(X^v)(w) \\ &= (dL_h)_w(X^v(w)) \\ &= (dL_h)_w\big((dL_w)_e(v)\big) \\ &= d(L_h \circ L_w)_e(v) \\ &= (dL_{hw})_e(v) \\ &\equiv X^v \circ L_h(w) \end{aligned}$$

So it is left-invariant.

Now define a map the other way:

$$\mathfrak{X}(G)^{L_g} \to T_eG$$

$$X \mapsto X_e$$

Exercise: Check that these two maps are inverses.

□

So far, we have only seen that the Lie algebra of a Lie group is a vector space. Later when we meet abstract Lie algebras, we will establish more structure on this vector space, so that the sentence "The Lie algebra of a Lie group is, in fact, a Lie algebra" is correct.

Definition: Given a matrix M, the exponential of the matrix is given by

$$exp(M) = \sum_{n=1}^{\infty} \frac{M^n}{n!}$$

The convergence of this series is a standard analysis exercise, which we leave for the reader.

Note $R_\theta = exp(\theta X)$, so $\frac{d}{d\theta}|_{\theta=0}$ goes from $G \to \mathfrak{g}$ and $exp : \mathfrak{g} \to G$. These maps are not necessarily inverses.

Functoriality: There is a category of Lie groups. The objects are Lie groups, and the morphisms are smooth group homomorphisms. This is a (not full) subcategory of the category of smooth manifolds[a] There is also a category of vector spaces, whose objects are vector spaces over $\mathbb{R}$ and morphisms are $\mathbb{R}$-linear maps. We now describe the "Lie functor" $Lie : LieGrp \to Vect_{\mathbb{R}}$. On objects, it sends $G \mapsto T_eG \equiv \mathfrak{g} \equiv Lie(G)$. Given a morphism $\varphi : G_1 \to G_2$, we can induce a map $\mathfrak{g}_1 \to \mathfrak{g}_2$ by

$$d\varphi_e : T_{e_1}G_1 \to T_{\varphi(e_1)}G_2 = T_{e_2}G_2$$

We can differentiate φ because it is required to be smooth, and $\varphi(e_1) = e_2$ because it is required to be a group homomorphism. That Lie respects compositions is exactly given by the chain rule, and clearly the identity morphism is sent to the identity morphism. Thus Lie is a functor.

TODO: is exp a functor? I think I read somewhere it can be regarded as a natural transformation between Lie and forget?

[a] I suppose you could also define it as a subcategory of Grp, but for some reason this way seems more natural.

Consider rotations in 3 dimensions now, corresponding to the Lie group $SO(3)$.

Matrix Lie groups: Generally, every matrix group which we will encounter in this class will be a Lie group.

Consider the set $M_n(\mathbb{R})$, the set of $n \times n$ matrices with entries in $\mathbb{R}$. There is an inclusion $M_n \to \mathbb{R}^{n^2}$, given by sending a matrix A_{ij} to $(a_{11}, a_{12}, a_{13}, \ldots, a_{nn})$. $\mathbb{R}^{n^2}$ has a trivial manifold structure, when equipped with the standard topology (open balls). The subset $GL(n, \mathbb{R}) \subset M_n(\mathbb{R}) \subset \mathbb{R}^{n^2}$ is an open subset: The map $det : \mathbb{R}^{n^2} \to \mathbb{R}$ is continuous: it is a polynomial in each coordinate. $GL(n, \mathbb{R})$ is the set $det^{-1}(\mathbb{R} \setminus 0)$, thus it is the preimage of an open set under a continuous function, and is thus continuous. So $GL(n, \mathbb{R})$ is an open subset of the smooth manifold $\mathbb{R}^{n^2}$, and thus inherits a smooth manifold structure. One can check it is also a Lie group: matrix multiplication is also polynomial. Inversion is a little trickier, but I believe one can cite the Cramer inverse formula. For the particular Lie groups such as $SL_n(\mathbb{R})$ and $O(n)$, we need to make other arguments.

For example, for $SL_n(\mathbb{R})$, we can employ the regular level set theorem, which says that the preimage of a regular value under a smooth map is a smooth manifold. Again we use the determinant map, but now consider the preimage of 1. So one must compute the derivative and actually show it is surjective. We leave this as an exercise.

Elements of $SO(3)$ look like

$$R_{x,\theta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

and there are similar formulas for rotations about y and z axes. These 3 generate the Lie group $SO(3)$. If we apply the Lie functor to $SO(3)$, we can see (by Taylor expanding), that the corresponding generators of the Lie algebra $\mathrm{Lie}(SO(3)) \equiv \mathfrak{so}(3)$, are given by

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

and similarly for Y and Z. In general, we often denote the Lie algebra of a Lie group as the same characters, but written lowercase in gothic font, as above.

As we saw in the definition, the Lie algebra is meant to be a sort of "zoomed in" approximation of the Lie group at the identity. But as we see, the operation on the Lie algebra,

which is a vector space, is always commutative. But the Lie group itself may not be. So we should introduce some kind of non-commutative structure to keep track of that fact. In this case, we define the commutator on $\mathfrak{so}(3)$,

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

where the product here is just given by matrix multiplication. Multiplying out, one can see that

$$[X, Y] = Z$$
$$[Y, Z] = X$$
$$[Z, X] = Y$$

Theorem (Baker-Campbell-Hausdorff):

$$exp(A)exp(B) = exp\Big(A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A[A, B]] + [B, [B, A]]) - \dots\Big)$$

□

So the expansion is given in terms of commutators. In particular, if A and B commute, then it reduces to the usual formula which we know for numbers. In particular, if we know all of the commutators in the Lie algebra, then in principle, we can work out the multiplciation in the group itself, because the exponentials are the Lie group elements[6].

[6]Modulo exp failing to be surjective, which happens sometimes.

CHAPTER 6

VI. Lie Theory

Lecture 6, Sept 14.

What is Lie($SO(3)$) as a vector space? Well we know it is supposed to consist of tangent vectors at the identity. Such a thing can be written in the form $1 + \epsilon X$, for $X \in M_n(\mathbb{R})$. But we only want to consider those directions which stay in $SO(3)$, in other words, we want that $\det(1 + \epsilon X) = 1$ and $(1 + \epsilon X)(1 + \epsilon X)^T = Id$. These conditions together imply that $X + X^T = 0$. In general, det 1 always implies traceless, so it is worth it to write out that one:

$$det(1 + \epsilon X) \approx 1 + TrX + O(\epsilon^2)$$

so that

$$det(1 + \epsilon X) = 1 \Rightarrow TrX = 0$$

So "special" matrix Lie groups always have Lie algebras which consist of traceless matrices, with some other conditions possibly. It just so happens that in this case, the "orthogonality" condition did not contribute any new constraints on the Lie algebra[1]

Definition: A Lie algebra is a vector space, $\mathfrak{g}$, equipped with a "Lie bracket" operation

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

which is anti-symmetric, $\mathbb{R}$-linear, and obeys the Jacobi Identity: for all $X, Y, Z \in \mathfrak{g}$

$$[[X, Y], Z] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

Example: One Lie group we will see often is $SU(2)$, the special unitary group of 2x2 matrices. It is easy to show that any such matrix in $SU(2)$ has the form

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

where $a, b \in \mathbb{C}$ such that $a^2 + b^2 = 1$. This is recognized as the unit sphere $S^3 \subset \mathbb{C}^2$.

[1] Actually he said something at this point which made me believe this is not a coincidence, but I didn't quite catch what he meant. Something like near the identity, orthogonality implies special? I'm not really sure.

Fact: $det(exp(A)) = exp(TrA)$. So for $A \in \mathfrak{su}(2)$, $exp(A) \in SU(2)$, then $Tr(A) = 2\pi ik$, for some $k \in \mathbb{Z}$. However, the trace function is continuous, so its image must be connected, and thus constant. But which k? Well we know $0 \in \mathfrak{su}(2)$, since it is a vector space. Thus $Tr0$ is in the image. But $Tr0 = 0$, so the trace is constant, and thus must be 0. So $\mathfrak{su}(2)$ has only traceless matrices[2]. To see what the unitarity condition implies, we must expand to first order in ϵ,

$$(1 + \epsilon X)(1 + \epsilon X)^T = Id \Rightarrow X = X^T$$

So $\mathfrak{su}(2)$ consists of traceless, symmetric matrices. A basis of such a space is given by

$$\hat{X} = \frac{1}{2i}\sigma_1, \quad \hat{Y} = \frac{1}{2i_2}, \quad \hat{Z} = \frac{1}{2i}\sigma_3$$

where

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

These σ_i are known as the Pauli spin matrices. One can calculate

$$[\hat{X}, \hat{Y}] = \hat{Z}$$
$$[\hat{Y}, \hat{Z}] = \hat{X}$$
$$[\hat{Z}, \hat{X}] = \hat{Y}$$

which we observe are the same relations as $\mathfrak{so}(3)$. This implies they are isomorphic, where an isomorphism of Lie algebras is just a vector space isomorphism which respects the Lie bracket.

But $SU(2) \ncong SO(3)$[3] One can see this by considering

$$exp(2\pi Z) = Id, \quad exp(2\pi \hat{Z}) = -Id$$

It turns out the center[4] of $SO(3) = \{1\}$, while the center of $SU(2)$ is $\{\pm 1\}$. And it turns out, this is the only obstruction to the isomorphism, in the sense that

$$SU(2) \cong SO(3)/\{\pm 1\}$$

So that $SU(2)$ is a double cover of $SO(3)$. If one knows some algebraic topology, this implies that $SO(3) \cong \mathbb{R}P^3$.

Now we move onto representation theory:

Definition: A representation of a group G is a group homomorphism $G \to End(V)$. V is often referred to as the representation space. The dimension of the representation is the

[2]Or one could apply the argument we made above, but this is actually a neat trick which I had not seen before.

[3]For example, they have different fundamental groups.

[4]The set of all elements which commute with all other elements

dimension of the representation space.

Example: Every group admits the trivial representation, sending $g \mapsto Id \in End(V)$, for all g.

Definition: A Lie algebra representation is a Lie algebra morphism $\mathfrak{g} \to \mathfrak{gl}(V)$, the Lie algebra of endomorphisms with commutator as the bracket.

One thing we would like to do is find a Lie algebra representation, then exponentiate it to get a Lie group representation.

Consider $SU(2)$ again. Choose a new basis:

$$\begin{aligned} H &= -2i\hat{Z} \\ U^+ &= i\hat{X} + \hat{Y} \\ U^- &= i\hat{X} - \hat{Y} \\ \Rightarrow [H, U^+] = 2U^\dagger, &\quad [H, U^-] = -2U^-, \quad [U^+, U^-] = H \end{aligned}$$

Suppose you have a representation $\rho : \mathfrak{su}(2) \to \mathfrak{gl}(\mathbb{C}^n)$, and suppose $v \in V$ is an eigenvector of H, with eigenvalue k. Denote the k eigenspace of H as H_k. Then applying the commutation relations,

$$\begin{aligned} H(U^+v) &= (U^+H + [H, U^+])v \\ &= U^+(H+2)v \\ &= (k+2)U^+v \\ &\Rightarrow U^+v \in H_{k+2} \end{aligned}$$

and similarly for U^-. In other words, U^+ sends k eigenvectors to $k+2$ eigenvectors (wrt H), and U^- drops the eigenvalue by 2. But we recall that eigenvectors belonging to distinct eigenvalues are linearly independent. Thus if we want this representation to be finite dimensional, then this process of increasing and decreasing eigenvalues must both terminate eventually, otherwise H has infinitely many distinct eigenvalues. Thus there must exist a v, eigenvalue of H, say of eigenvalue w, such that

$$U^+v = 0$$

and there must be an $n \in \mathbb{N}$ such that

$$(U^-)^n v = 0$$

with $(U^-)^{n-1}v \neq 0$. So we have a collection of vectors,

$$\left\{v, U^-v, (U^-)^2v, \ldots, (U^-)^{n-1}v\right\} \equiv W$$

which are all linearly independent. Thus they span an n-dimensional subspace of V. Further, this subspace is invariant under the action of $\mathfrak{su}(2)$, in the sense that, for any

$X \in \mathfrak{su}(2)$ and $w \in W$, $Xw \in W$. This can be checked by just writing w as a combination of the spanning vectors, and applying H, U^+ and U^-. Thus W is an invariant subspace of the representation space V.

Definition: A representation $\mathfrak{g} \to \mathfrak{gl}(V)$ is irreducible if it contains no non-trivial invariant subspaces (the trivial subspaces are 0 and V). In particular, if there is a decomposition $V = V_1 \oplus V_2$ into invariant subspaces, then either V_1 or V_2 equals V, and thus the other is 0.

So we have been considering an arbitrary finite dimensional representation of $SU(2)$, and found that it has an n-dimensional invariant subspace. If we want to consider *irreducible* such representations, then we must have the W above equal to V (it can't be zero because it contains v). So V has this basis given by the highest weight vector v, and is n-dimensional.

As a result, we have identified what every finite dimensional irrep must look like. The full situation is much better, though:

Theorem (Classificiation of Reps of $\mathfrak{su}(2)$): *For every $n \in \mathbb{N}$, there exists a unique irreducible representation of $\mathfrak{su}(2)$ with dimension n, of the form as above.*

We didn't use the fact that this is a Lie algebra representation yet. If one applies the commutation relations from $\mathfrak{su}(2)$, we get that for every m,

$$U^+(U^-)^m(v) = m(w - m + 1)(U^-)^{m-1}v$$

In particular,

$$U^+(U^-)^n(v) = n(w - n + 1)(U^-)^{n-1}v$$

But we know the left hand side must be zero, which implies $n = w + 1$. So $w = n - 1$ is also an integer. Sometimes people classify the irreps of $\mathfrak{su}(2)$ by $w = n - 1$, the eigenvalue of the highest weight vector, sometimes referred to as the irrep of highest weight $n - 1$. This often leads to confusion.

So the possible eigenvalues are[5]

$$\Big\{w, w-2, w-4, \ldots, 2-w, -w\Big\}$$

Every Lie algebra comes with a canonical representation, the adjoint representation:

Definition: The adjoint representation of a Lie algebra $\mathfrak{g}$ is defined as the map

$$ad : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

[5]I didn't catch the argument for why must end at $-w$.

$$X \mapsto ad_X \equiv [X, -]$$

i.e. $ad_X(Y) = [X, Y]$. This is clearly an $\mathbb{R}$-linear map, and the fact that this action respects the Lie bracket structure is exactly equivalent to the Jacobi identity[6]

Given two representations, V_1, V_2, one can construct new representations $V_1 \oplus V_2$ and $V_1 \otimes V_2$, where the group (or Lie algebra) acts on each component.

[6]It is often discussed that the Lie algebra rep condition follows from Jacobi, but the other implication is also true. This is one of the better ways of internalizing the Jacobi identity, I feel, rather than a silly relation among brackets. We will see another equivalent condition later as well.

CHAPTER 7

VII. More Lie Theory

Lecture 7, Sept 19.