Representation Theory And Quantum Mechanics

Noah Miller, January 2018

Preface

I got the urge to review representation theory while writing a set of notes on quantum field theory. My eventual goal was to develop the theory of the Poincaré group for quantum field theory, but the preamble just kept growing and growing! What I am left with instead is a unique set of notes that explain the relationship between representation theory and *non-relativistic* quantum mechanics.

A few years ago, I took Peter Woit's class on representation theory and quantum mechanics. The notes for that class have now been published as a book (which I am acknowledged in!). I used his book extensively while writing these notes, which condense approximately the first half of his book.

What is unique about these notes is that they are constantly searching for the "moral" of the math. Why are certain relationships and results important? In a subject as deceptively simple as this one, stepping back and untangling what you've done can often be the hardest part. That was certainly the case for me.

Perhaps one frustrating aspect of these notes is that they leap from topic to topic, explaining them tersely before moving on to the next one unceremoniously. The topic with the least explicit foreshadowing is the relationship between Hamiltonian mechanics, the moment map, and quantization, which is the last topic of these notes. The story of the moment map is perhaps my favorite one here, if only because it is poorly explained everywhere else. The moment map gives a much richer connection between symmetries and conserved quantities than Noether's theorem, and should be better known by physics students. Having said that, I truly love all of the material in these notes. It's nothing but net.

I must thank **Naomi Sweeting** for her extensive help with these notes and **Theo Coyne** for his late-night photography.

The following books were consulted in varying degrees during the writing of these notes:

- *Quantum Theory, Groups and Representations: An Introduction* - Peter Woit
- *Algebraic Topology* - Allen Hatcher
- *Gauge Fields, Knots and Gravity* - John Baez & Javier P. Muniain
- *The Quantum Theory of Fields, Volume 1: Foundations* - Steven Weinberg

Contents

1 Group Representations: Definitions **3**

2 Schur's Lemma **4**
2.1 The Statement . . . 4
2.2 Proof of Part 1 . . . 5
2.3 Proof of Part 2 . . . 5
2.4 Corollary 1 . . . 6
2.5 Corollary 2 . . . 6
2.6 Why is Schur's Lemma, and Representation Theory, so important? . . 7

3 Spherical Harmonics as Representations of $SO(3)$ **8**

4 Unitary Representations **10**

5 All Representations Break Up Into a Direct Sum of Irreducible Representations* **11**

6 A Brief Safari of The Hydrogen Atom **11**

7 Lie Algebras **13**

8 The Lie Algebras of $U(n), SL(n, \mathbb{C})$, and $SU(n)$ **16**

9 Lie Algebra Representations **18**

10 Classifying the Irreducible Representations of $U(1)$ **19**

11 Lie Algebra Complexifications **20**

12 Classifying the Irreducible Representations of $SU(2)$ **22**

13 The Spin 1 Representation of $SU(2)$ **26**

14 An Explicit Construction of Unitary $SU(2)$ Representations **26**

15 The Odd Couple: $SU(2)$ and $SO(3)$ **28**

16 What's The Deal With Spin? **34**

17 The Adjoint Representation And The Power Of Magical Thinking **36**

18 Operators That Generate Representations of $U(1)$ Have Quantized Eigenvalues **43**

19 Functions on Phase Space Comprise a Lie Algebra **47**

20 The Moment Map: Lie Algebra $\rightarrow$ Conserved Quantities **49**

21 Quantization is a Lie Algebra Representation **53**

22 Quantizing a Group Action **54**

23 Symplectomorphisms and Degree 2 Polynomials **58**

1 Group Representations: Definitions

A **group** G is a set with a composition operation $* : G \times G \to G$ that satisfies the following three axioms:

- There is an identity element $1 \in G$ such that, for all $g \in G$, $1 * g = g * 1 = g$.
- For each $g \in G$, there is an inverse element g^{-1} such that $g * g^{-1} = g^{-1} * g = 1$.
- $*$ is associative. That is, $a * (b * c) = (a * b) * c$ for all $a, b, c \in G$.

It's not hard to prove that the identity element is unique and that g^{-1} for each g. (Also, not all of the above axioms are necessary. Some portions of the axioms can be derived from the others. Can you figure out which?) Usually we will drop the composition symbol $*$.

A group G is **commutative** or **abelian** if $g * h = h * g$ for all $g, h \in G$.

An element $g \in G$ is called **central** if it commutes with every element in G.

A group G has a **left action** on a set X if $g \cdot x \in X$ for all $g \in G$ and $x \in X$ and the following axioms are satisfied:

- $1 \cdot x = x$
- $g \cdot (h \cdot x) = (g * h) \cdot x$

(Usually we drop the $\cdot$.) By contrast, G is said to have a **right action** if we have

- $g \cdot (h \cdot x) = (h * g) \cdot x$

instead, but we'll never talk about right actions ever again!

Note that if you have a group G with a (left) action on a set X, then you also have a (left) group action on the set of functions from X to some other set Y, where the group action is defined as follows, where $f : X \to Y$ is our function:

$$(g \cdot f)(x) = f(g^{-1} \cdot x)$$

Let's confirm this that this satisfies the two axioms.

- $(1 \cdot f)(x) = f(1^{-1} \cdot x) = f(x)$
- $\big(g \cdot (h \cdot f)\big)(x) = (h \cdot f)(g^{-1} \cdot x) = f\big(h^{-1} \cdot (g^{-1} \cdot x)\big) = f\big((g*h)^{-1} \cdot x\big) = \big((g*h) \cdot f\big)(x)$

Great! Note that the inverse in the definition of the group action is vital.

For an n-dimensional vector space V, the **general linear group** $GL(V)$ is the set of all $n \times n$ invertible matrices that map $V \to V$. Note that this is a group, if the name didn't make that clear. Vitally, $GL(V)$ acts on V in the obvious way. If V is the n dimension vector space over the field F, then this is often written as $GL(n, F)$. We'll only deal with the fields $\mathbb{R}$ and $\mathbb{C}$.

For the groups G and H, a **group homomorphism** $\pi : G \to H$ is a map that respects the group structure in the following way:

$$\pi(g_1 * g_2) = \pi(g_1) * \pi(g_2) \text{ for all } g_1, g_2 \in G$$

From this you can derive that

$$\pi(1_G) = 1_H$$

and

$$\pi(g^{-1}) = \pi(g)^{-1} \text{ for all } g \in G.$$

A **Lie group** is a group that is also a manifold where the action of a group element on the group itself is a smooth map.

A group homomorphism π is a **representation** of a group G if its domain is $GL(V)$ for some vector space V.

$$\pi : G \to GL(V)$$

Note that a representation π depends on the vector space V being considered. Two representations are not the same if they map into different general linear group! Sometimes we will distinguish representations by referring to the *vector spaces* they act on. This is often confusing for newcomers! This is only done when the definition homomorphism π itself is obvious. If people (including me) ever refer to representations by referring to a vector space, and you become confused what the representation actually *is*, just try and write down what the homomorphism $\pi \to GL(V)$ is for yourself.

Two representations $\pi_1 : G \to GL(V)$ and $\pi_2 : G \to GL(V)$ are **equivalent** if there is a matrix $A : V \to V$ such that

$$\pi_1(g) = A^{-1}\pi_2(g)A$$

for all $g \in G$.

Note that if we have two representations $\pi_1 : G \to V_1$ and $\pi_2 : G \to V_2$ (assuming V_1 and V_2 are vector spaces over the same field) we can define a representation $\pi_1 \oplus \pi_2 \to GL(V_1 \oplus V_2)$ as

$$\big(\pi_1 \oplus \pi_2\big)(g) = \pi_1(g) \oplus \pi_2(g) = \begin{pmatrix} \pi_1(g) & \mathbf{0} \\ \mathbf{0} & \pi_2(g) \end{pmatrix}.$$

Recall that $V_1 \oplus V_2$ is the **direct sum** of V_1 and V_2. In the above example, π_1 and π_2 are both **sub-representations** of π. We would say that any representation equivalent to $\pi_1 \oplus \pi_2$ also has sub-representations, even if they can't be expressed in the block diagonal form above.

An **irreducible representation** (often called an "irrep") is a representation with no sub-representations (except for the trivial one and itself). All the representations we will study will break up into a direct sum of irreducible representations. This isn't true of all Lie groups, just the easy ones that we physicists most like to study.

2 Schur's Lemma

2.1 The Statement

Let π_1 and π_2 be irreducible representations of the group G and V_1 and V_2 are complex vector spaces.

$$\pi_1 : G \to GL(V_1)$$

$$\pi_2 : G \to GL(V_2)$$

Suppose the linear transformation $S : V_1 \to V_2$ is an "intertwining operator" between these two representations. That is, suppose

$$S\pi_1(g) = \pi_2(g)S \text{ for all } g \in G.$$

Schur's Lemma states that

1. If π_1 and π_2 are inequivalent representations, then $S = \mathbf{0}$.

2. If $\pi_1 = \pi_2$ (implying $V_1 = V_2$) then $S = \lambda\mathbf{1}$ for some $\lambda \in \mathbb{C}$.

2.2 Proof of Part 1

For all $g \in G$,

$$\text{for all } v_1 \in \ker S \qquad S\big(\pi_1(g)v_1\big) = \pi_2(g)S(v_1) = \pi_2(g)0 = 0$$
$$\text{so } \pi_1(g)v_1 \in \ker S$$

and

$$\text{for all } v_2 \in \operatorname{Im} S \text{ there exists a } v_1 \in V_1 \text{ such that } v_2 = S(v_1)$$
$$\text{so } \pi_2(g)v_2 = \pi_2(g)S(v_1) = S\big(\pi_1(g)v_1\big) \in \operatorname{Im} S$$

Therefore, for all $g \in G$, $\pi_1(g)\big(\ker S\big) = \ker S$ and $\pi_2(g)\big(\operatorname{Im} S\big) = \operatorname{Im} S$. This proves that $\ker S$ is a sub-representation of π_1 and $\operatorname{Im} S$ is a sub-representation of π_2.

As π_1 and π_2 are both irreducible representations, the only sub-representations they have are the trivial one and themselves. Therefore we have two possibilities.

1. $\ker S = \mathbf{0}$ and $\operatorname{Im} S = V_2$. This implies that S is invertible, meaning $S\pi_1(g)S^{-1} = \pi_2(g)$ for all $g \in G$, which is a contradiction as π_1 and π_2 are inequivalent.

2. $\ker S = V_1$ and $\operatorname{Im} S = \mathbf{0}$. This implies $S = \mathbf{0}$, completing the proof. □

2.3 Proof of Part 2

We have the irreducible representation

$$\pi : G \to GL(V)$$

for a complex vector space V and an operator S that commutes with every element of the representation:

$$S\pi(g) = \pi(g)S \text{ for all } g \in G$$

By the fundamental theorem of algebra, there is a $\lambda \in \mathbb{C}$ such that $\det(S - \lambda\mathbf{1}) = 0$. As

$$\big(S - \lambda\mathbf{1}\big)\pi(g)v = \pi(g)\big(S - \lambda\mathbf{1}\big)v$$

if $v \in \ker\big(S - \lambda\mathbf{1}\big)$, then $\pi(g)v \in \ker\big(S - \lambda\mathbf{1}\big)$:

$$v \in \ker\big(S-\lambda\mathbf{1}\big) \quad\Longrightarrow\quad \big(S-\lambda\mathbf{1}\big)\pi(g)v = \pi(g)\big(S-\lambda\mathbf{1}\big)v = 0 \quad\Longrightarrow\quad \pi(g)v \in \ker\big(S-\lambda\mathbf{1}\big)$$

Therefore $\ker\big(S - \lambda\mathbf{1}\big)$ is a sub-representation of π_1. As π_1 is an irreducible representation and $\ker\big(S - \lambda\mathbf{1}\big)$ is at least 1-dimensional, $\ker\big(S - \lambda\mathbf{1}\big)$ must be all of V. Therefore $S = \lambda\mathbf{1}$. □

2.4 Corollary 1

If G is commutative then all of its irreducible representations over complex vector spaces are 1-dimensional.

Proof: If π is a representation of G, then

$$\pi(g_1)\pi(g_2) = \pi(g_2)\pi(g_1)$$

for all $g_1, g_2 \in G$. By Schur's lemma, all $\pi(g)$ must be scalar matrices.

$$\pi(g) = \lambda_g \mathbf{1}$$

In order for such a π to be to be irreducible it must be one-dimensional. $\square$

2.5 Corollary 2

Say the representation $\pi : G \to GL(V)$, where V is a complex vector space, can be broken up into the irreducible representations $\pi_k : G \to GL(V_k)$ as follows:

$$\pi(g) = \begin{pmatrix} \pi_1(g) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \pi_2(g) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \pi_n(g) \end{pmatrix}$$

where all representations π_k are inequivalent to each other. If an operator $H : V \to V$ commutes with the whole representation

$$\pi(g)H = H\pi(g) \text{ for all } g \in G$$

then H is of the form

$$H = \begin{pmatrix} E_1\mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & E_2\mathbf{1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & E_n\mathbf{1} \end{pmatrix}$$

for some constants E_k.

Proof: Define the operators $P_k : V \to V$ to be the projection operators onto the subspace $V_k \subset V$ of the k^{th} irreducible representation. Note that $\big(P_k H P_l\big)$ commutes with all elements of the representation.

$$\text{For all } g \in G, \qquad \big(P_k H P_l\big)\pi(g) = P_k H \pi(g) P_l = P_k \pi(g) H P_l = \pi(g)\big(P_k H P_l\big).$$

Note that the map $\big(P_k H P_l\big)$ can also be regarded a map from $V_l \to V_k$, and

$$\text{for all } g \in G, \qquad \big(P_k H P_l\big)\pi_l(g) = \pi_k(g)\big(P_k H P_l\big).$$

Therefore, By Schur's Lemma, if π_k is inequivalent to π_l then $\big(P_k H P_l\big) = \mathbf{0}$. If π_k is equivalent to π_l then $\big(P_k H P_l\big)$ is a constant multiple of the identity. $\square$

Note: In general, a given representation may have repeated irreducible representations. Say, for example, π is an irreducible representation. Then the representation $\pi \oplus \pi$, defined by

$$\big(\pi \oplus \pi\big)(g) = \begin{pmatrix} \pi(g) & \mathbf{0} \\ \mathbf{0} & \pi(g) \end{pmatrix},$$

cannot be decomposable as a direct sum of inequivalent irreps. Therefore, for an operator H that commutes with every element $\big(\pi \oplus \pi\big)(g)$, our proof above does not apply. However, it does show that H must be of the following form, for some constants A, B, C, D:

$$H = \begin{pmatrix} A\mathbf{1} & B\mathbf{1} \\ C\mathbf{1} & D\mathbf{1} \end{pmatrix}.$$

2.6 Why is Schur's Lemma, and Representation Theory, so important?

Let me just say if you don't understand the proof of Schur's Lemma you have no hope of understanding representation theory. Set aside a few hours to get it through your thick skull. (Believe me, it's tough for ~~everyone~~ every physicist their first time.)

Representation theory is important in physics for a billion reasons, but here is one: Hilbert spaces are complex vector spaces, so any group action on a Hilbert space will necessarily be a group representation. Therefore, if we want to understand how groups can act on quantum mechanical state spaces, we must understand representation theory. This shouldn't be very surprising. Quantum mechanics is really just glorified linear algebra, and representation theory is all about using linear algebra to study groups.

Schur's Lemma's is important for a few reasons. For one, it shows how inequivalent irreps are fundamentally different from each other. There's no way to "intertwine" them without doing it trivially! (Question: the proof of Schur's Lemma relied on the fact that the vector space was complex in order to find an eigenvalue of S using the fundamental theorem of algebra. So wouldn't it also work on real vector spaces of odd dimension? No one ever mentions this...)

More relevantly, the importance of Schur's Lemma to physics can be seen most easily through Corollary 2 above. As you probably guessed, it was no coincidence that that my operator "H" had eigenvalues "E." *If you have a group representation acting on your Hilbert space that commutes with your Hamiltonian, the energy eigenspaces will be irreducible representations* (except for equivalent irreps, and usually even then). This has two important consequences (e.f.e.i.a.u.e.t):

- A state that lives in an irrep will stay there under time evolution.
- Acting on an energy eigenstate by your group will not change its energy.

But even if a representation doesn't commute with the Hamiltonian, the mere *presence* of a representation on a state space is already pretty interesting. This is because a representation can be broken up into irreducible representations. This means that our Hilbert space can be broken up into a collection of irreducible representations, even if they interact in a non-trivial way with the Hamiltonian.

I must reiterate that the importance of representation theory in physics cannot be summed up succinctly. For example, in QFT, the Poincaré group does not commute with the Hamiltonian, as a boost will in general change the energy of a state. Even so, the Poincaré group is still very important to study!

There is one last consequence that comes from the Hamiltonian commuting with a representation. Namely, the Lie algebra elements that "generate" the representation will be conserved quantities under time evolution. (I haven't told you yet what a Lie algebra is yet, but be patient!)

3 Spherical Harmonics as Representations of $SO(3)$

The group $SO(3)$ is the group of of real 3×3 special orthogonal matrices, A.K.A., the group of rotations. This is a group with a natural action on $\mathbb{R}^3$, but let's study it's representations on complex vectors spaces. We will denote elements of $SO(3)$ by "R."

So: what are the irreducible representations π of $SO(3)$?

$$\pi : SO(3) \to GL(\mathbb{C}^n)$$

Well, obviously there's the trivial representation $\pi(R) = 1$.

$$\pi : SO(3) \to GL(\mathbb{C}^1)$$

This is a 1-dimensional representation. There's also obviously the "defining" representation of $SO(3)$. This is just the one where the matrix $\pi(R)$ is just the same as R.

$$\pi : SO(3) \to GL(\mathbb{C}^3)$$

Are there any others?

You'll have to take my word for it now, but all irreps of $SO(3)$ have odd dimension. Furthermore, any two irreps of $SO(3)$ with the same dimension are equivalent. *Furthermore*, there is one irrep of $SO(3)$ for each odd dimension. So there is one irrep of $SO(3)$ with dimension 1, one with dimension 3, one with dimension 5, etc. As is standard we refer to the different irreducible representations of $SO(3)$ with a non-negative integer "l." The dimension of the l representation is $2l + 1$. So the trivial representation is the "$l = 0$" representation, the defining representation is the "$l = 1$" representation, and so on.

Perhaps the most natural space $SO(3)$ acts on is S^2, the 2-sphere. (Recall that a 2-sphere lives in $\mathbb{R}^3$ as the set of points (x, y, z) where $x^2 + y^2 + z^2 = 1$.) Namely, for any $R \in SO(3)$ and $\vec{x} \in S^2$, $R\vec{x}$ is just action of the matrix R on the vector $\vec{x}$.

L^2 complex functions of points in S^2, i.e. functions of the form

$$f : S^2 \to \mathbb{C}$$

comprise the set "$L^2(S^2)$." In other words, $f \in L^2(S^2)$. Note that $L^2(S^2)$ is an infinite dimensional complex vector space. (Even though all of our definitions and proofs so far have only applied for finite dimensional vector spaces, we're physicists so we don't care.)

There is a natural representation of $SO(3)$ that acts on the vector space $L^2(S^2)$:

$$\pi : SO(3) \to GL\big(L^2(S^2)\big)$$

$$\big(\pi(R)f\big)(\vec{x}) = f(R^{-1}\vec{x})$$

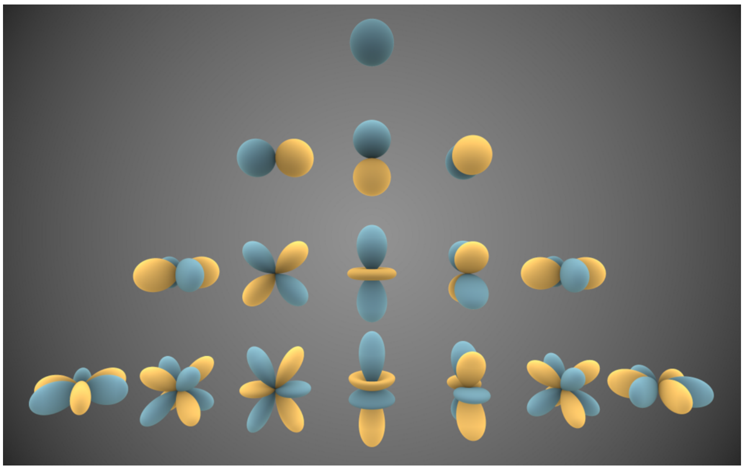

Figure 1: Taken from the Wikipedia article *Spherical harmonics*: "Visual representations of the first few real spherical harmonics. Blue portions represent regions where the function is positive, and yellow portions represent where it is negative. The distance of the surface from the origin indicates the value of $Y_l^m(\theta, \phi)$ in the angular direction (θ, ϕ)."

How can we decompose this complex representation of $SO(3)$ into a direct sum of irreducible representations? It turns out that every irreducible representation of $SO(3)$ is present in our representation exactly *once*!

$$\pi = \bigoplus_{l=0}^{\infty} \pi_l$$

Wow! What are they?

Let's start with the $l = 0$ representation. What could it be? We're looking for functions on the 2-sphere that don't change after being rotated. Clearly, the only functions that satisfy this are the constant functions. And indeed, the constant functions on S^2 are a 1 dimensional vector space! (1 *complex* dimension, that is.)

Let me just spoil the answer: the irreducible sub-representations of π are spherical harmonics!

Spherical harmonics are functions that comprise a basis of $L^2(S^2)$. Each function $f \in L^2(S^2)$ can be written as follows, for some choice of constants $a_{ml} \in \mathbb{C}$:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{m,l} Y_l^m(\theta, \phi)$$

(Points in S^2 are often denoted by their angular coordinate (θ, ϕ).) Spherical harmonics are labeled by two numbers: l and m. l appropriately specifies which subrepresentation of π the spherical harmonic is in. m further specifies which basis element of representation l the function is. So for each l, there are $2l + 1$ different m's, ranging from $-l$ to l. Sperical Harmonics are often referred to as $Y_l^m(\theta\phi)$. Figure 1 shows these functions (chosen in a particular way so that they're all real functions instead of complex).

Let me now explain what I mean when I say that they "are irreducible representations of $SO(3)$." For a particular l, the spherical harmonics with that particular l comprise a $2l + 1$ dimensional vector space. This vector space is a subspace of

$L^2(S^2)$. Take any random vector from this subspace. If $l = 2$, for example, such a function could be expressed as follows:

$$f(\vec{x}) = a_{-2,2} Y_2^{-2}(\vec{x}) + a_{-1,2} Y_2^{-1}(\vec{x}) + a_{0,2} Y_2^{0}(\vec{x}) + a_{1,2} Y_2^{1}(\vec{x}) + a_{2,2} Y_2^{2}(\vec{x})$$

f can be expressed using the vector $\vec{a} = (a_{-2,2}, a_{-1,2}, a_{0,2}, a_{1,2}, a_{2,2})$.

Now, if we operate on f for any $R \in SO(3)$, the resulting function will *still* be a linear combination of the 5 different $l = 2$ spherical harmonics!

$$\begin{aligned}
\big(\pi(R)f\big)(\vec{x}) &= f(R^{-1}x) \\
&= a_{-2,2} Y_2^{-2}(R^{-1}\vec{x}) + a_{-1,2} Y_2^{-1}(R^{-1}\vec{x}) + a_{0,2} Y_2^{0}(R^{-1}\vec{x}) + a_{1,2} Y_2^{1}(R^{-1}\vec{x}) + a_{2,2} Y_2^{2}(R^{-1}\vec{x}) \\
&= b_{-2,2} Y_2^{-2}(\vec{x}) + b_{-1,2} Y_2^{-1}(\vec{x}) + b_{0,2} Y_2^{0}(\vec{x}) + b_{1,2} Y_2^{1}(\vec{x}) + b_{2,2} Y_2^{2}(\vec{x})
\end{aligned}$$

This is extremely non-trivial behavior! If we think of f as the 5-dimensional vector $\vec{a}$ and $\pi(R)f$ as the 5-dimensional vector $\vec{b}$, then we can directly see how our representation acts on our function with the $l = 2$ representation.

$$\pi_2(R)\vec{a} = \vec{b}$$

If we then use our spherical harmonics as a basis of $L^2(S^2)$, with the subspace V_l defined by

$$V_l = \operatorname{span}\{Y_l^m\}_{m=-l,\ldots,l}$$

then our representation will be in block diagonal form, with each matrix $\pi_l(R)$ a $2l + 1$ dimensional matrix!

$$\pi(R) = \begin{pmatrix} \pi_0(R) & \mathbf{0} & \mathbf{0} & \ldots \\ \mathbf{0} & \pi_1(R) & \mathbf{0} & \ldots \\ \mathbf{0} & \mathbf{0} & \pi_2(R) & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

4 Unitary Representations

A linear transformation U is **unitary** if $U^\dagger = U^{-1}$. Unitary transformations are important in quantum mechanics because they don't affect the inner product of states. That is, the inner product between the state $U|\psi\rangle$ and $U|\phi\rangle$ is the same as the inner product between the state $|\psi\rangle$ and the state $|\phi\rangle$:

$$\langle\phi| U^\dagger U |\psi\rangle = \langle\phi| U^{-1} U |\psi\rangle = \langle\phi|\psi\rangle .$$

Because unitary transformations are so important in quantum mechanics, we are usually interested in representations that are unitary. A representation π is a **unitary representation** if all of the matrices $\pi(g)$ are unitary. In other words, π must be a map

$$\pi : G \to U(n)$$

where $U(n)$ is the group of $n \times n$ unitary matrices. Often, such representations are designated by the symbol "U" instead of our "π."

5 All Representations Break Up Into a Direct Sum of Irreducible Representations*

You may be wondering if every single representation decomposes into a unique direct sum of irreducible representations. Sadly, this is not the case.

For finite groups, finite dimensional representations always decompose into irreps. This is called "Maschke's Theorem" and isn't so hard to prove. But we're not interested in finite groups. We're interested in Lie groups!

As Lie Groups are manifolds, and therefore analytic objects, the situation gets more complicated. However, it turns out that the representation theory of a Lie group G is especially nice if G is compact. It turns out that all *unitary* representations of a compact Lie group break up into a direct sum of irreps. This is called the "Peter-Weyl theorem" and is beyond the scope of my brain.

Another useful thing to know: If a Lie group has a "semisimple" Lie Algebra (which we will discuss shortly) then all representations decompose into irreducible representations.

We will just assume in these notes that all of our Lie group representations will break up into a direct sum of irreducible representations. To actually show this, however, takes a bit of work!

Let me leave you with the canonical example of what can happen when a Lie group isn't compact. Take $\mathbb{R}$, for example, where the group operation is addition. We then have the following representation:

$$\pi(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

This is a homomorphism because

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+y \\ 0 & 1 \end{pmatrix}.$$

The vector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ transforms under this group action as the trivial representation, and yet these matrices are not diagonalizable. Therefore this representation does not split up into a direct sum of irreducible representations.

6 A Brief Safari of The Hydrogen Atom

The Hamiltonian of a (spinless) electron in the Coulomb potential is

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r}.$$

The Hilbert space of possible electron states is just functions from $\mathbb{R}^3 \to \mathbb{C}$.

$$\mathcal{H} = L^2(\mathbb{R}^3)$$

The group $SO(3)$ is unitarily represented on $\mathcal{H}$, by $U : SO(3) \to GL(\mathcal{H})$. The group action is the natural one. For any $R \in SO(3)$,

$$\big(U(R)\psi\big)(\vec{x}) = \psi(R^{-1}\vec{x}).$$

As the Hamiltonian is rotationally invariant, conjugation by $U(R)$ does not affect $\hat{H}$.

$$U(R)^{-1}\hat{H}U(R) = \hat{H}$$

Therefore $[\hat{H}, U(R)] = 0$ for all $R \in SO(3)$. From our discussion of Schur's Lemma, we know that irreducible sub-representations of U ought to be energy eigenspaces.

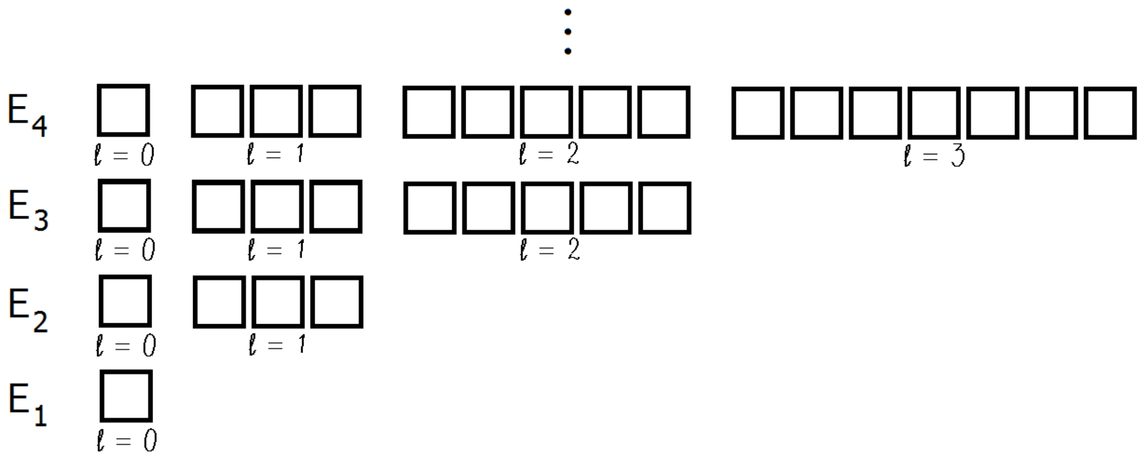

Figure 2: A schematic of the bound states of the Hydrogen atom. The energy levels are labeled by n, with $E_n = E_1/n^2$. The n^{th} energy level contains n different irreducible representations of $SO(3)$, $l = 0$ through $l = n - 1$. Each irreducible representation is spanned by the spherical harmonics multiplied by a radial function.

These eigenspaces are represented in Figure 2. A basis of energy eigenstates is indexed by three "quantum numbers." n, the energy level, along with l and m. We can see that the n^{th} energy eigenspace contains n different irreps of $SO(3)$. Wave functions are given by the appropriate spherical Harmonics multiplied by a radial function.

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_l^m(\theta, \phi)$$

The radial part does not depend on m. This is because any wave function that is a sum of ψ_{nlm}'s for a certain n and l, once rotated, must remain a sum of ψ_{nlm}'s for that n and l. The fact that the angular part is given by spherical harmonics is a direct consequence of the fact that our Hamiltonian commutes with our representation of $SO(3)$.

Note that there is a surprising amount of degeneracy in our energy eigenspaces. We expect some degeneracy from Schur's Lemma. That is, we expect for each energy eigenspace to contain at least one irrep of $SO(3)$. However, we can see that each energy eigenspace contains multiple irreps. This is degeneracy is more than we would expect from $\hat{H}$ commuting with our representation.

Where does this extra degeneracy come from? We've already noted that *some* degeneracy comes from the $SO(3)$ symmetry. We should therefore guess that the *extra* degeneracy comes from another, previously unidentified symmetry.

This is indeed the case. There is an extra hidden, completely non-obvious, $SO(4)$ symmetry hidden in our unassuming Hamiltonian. (This symmetry is even present in the classical Coulomb potential– Runge-Lenz vector, anyone?) I won't discuss this strange hidden symmetry, but Woit's book talks about it a bit. He even sketches how the energies can be exactly solved using just representation theory.

Most quantum mechanics textbooks solve for the definite energy wave-functions by tediously unwrapping differential equations. However, by examining the problem from the standpoint of representation theory, we can see that many properties of the Hydrogen atom arise directly from symmetries in a natural way.

Interestingly, the field of chemistry is largely governed by

1. the shape and energy of orbitals, which dictate how atoms like to bond
2. the fact that electrons are spin 1/2 particles which obey the Pauli Exclusion Principle.

We've already shed some light on how the first factor is a direct consequence of representation theory. It turns out that the second one is too! Apparently, it's not wrong to say that representation theory is the branch of math with the largest role to play in the chemical make-up of our world! (Shouldn't we tell the chemists?)

7 Lie Algebras

As we said before, a Lie group is just a group that is also a manifold. Whenever we think about points in a manifold, it is very natural to wonder about the tangent spaces of those points. Furthermore, the only truly distinguished point in a group is the identity element. Therefore, it is the mere *definition* of a Lie group that compels us to study the "tangent space at the identity."

Mathematicians have many ways of formulating the "tangent space," but because we're physicists we should just pick the easiest one. If we embed our Manifold M in some ambient space $\mathbb{R}^n$, then the tangent space at a point x, called T_xM, is the plane that best approximates the manifold near x. Tiny variations in x will be elements of the tangent space to the first order in the variations. For this reason, sometimes we associate the tangent space with the infinitesimal neighborhood of x, even though the tangent space is really a macroscopic thing.

Another way to think about the tangent space (still imagining it embedded in a larger ambient space) is as the vector space of velocity vectors of paths passing through x. This is depicted in Figure 3.

For a Lie Group G, we will define its Lie algebra $\mathfrak{g}$ as T_1G, the tangent space at the identity. (Often Lie algebras are denoted with the lowercase "fraktur" symbol of the group.) Therefore, we can see that, first and foremost, Lie algebras are vector spaces. A real linear combination of elements in $\mathfrak{g}$ is also in $\mathfrak{g}$.

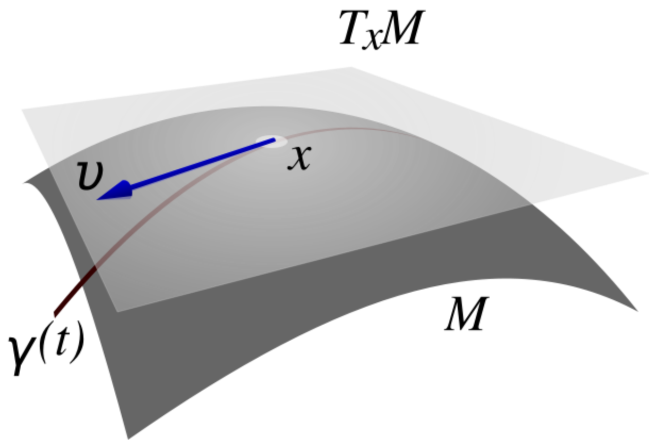

Figure 3: Image taken from the Wikipedia article *Tangent space*. For a manifold M the tangent space at the point $x \in M$ is denoted by $T_x M$. For a curve $\gamma(t)$ that passes through x at $t = 0$, the velocity vector $v = \frac{\mathrm{d}}{\mathrm{d}t}\, \gamma(t)\big|_{t=0}$ will be an element of $T_x M$.

As we are physicists, we will only concern ourselves with Lie algebras that are also matrix groups. Matrix groups have a natural embedding into ambient space. Namely, the n^2 numbers that label the cells in a matrix can just be taken to be a coordinate! So, for example, $GL(n, \mathbb{R}) \subset \mathbb{R}^{n^2}$ and $GL(n, \mathbb{C}) \subset \mathbb{R}^{2n^2}$.

The analog of the addition of elements of $\mathfrak{g}$ is the multiplication of elements of G. This can be seen because, for small ϵ and $X \in \mathfrak{g}$, we have a group element of the form $\mathbf{1} + \epsilon X + \mathcal{O}(\epsilon^2) \in G$. The composition of two such group matrices amounts to adding together the Lie algebra elements.

$$\big(\mathbf{1} + \epsilon X + \mathcal{O}(\epsilon^2)\big)\big(\mathbf{1} + \epsilon Y + \mathcal{O}(\epsilon^2)\big) = \mathbf{1} + \epsilon(X + Y) + \mathcal{O}(\epsilon^2)$$

However, because G is a group, the tangent space $T_{\mathbf{1}}G$ has many more interesting properties than just being a (real) vector space. Namely, for any $X \in \mathfrak{g}$, we have

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k \in G.$$

It is easy to see why this is. The following matrix is an element of G for any N (assuming N is larger than some unimportant lower bound):

$$\mathbf{1} + \tfrac{1}{N} X + \mathcal{O}\big(\tfrac{1}{N^2}\big) \in G.$$

The proof follows by haiku:

$$\begin{aligned} \lim_{N\to\infty} \Big(\mathbf{1} + \tfrac{1}{N} X + \mathcal{O}\big(\tfrac{1}{N^2}\big)\Big)^N &\in G \\ = \lim_{N\to\infty} \Big(\mathbf{1} + \tfrac{1}{N} X\Big)^N &\in G \\ = e^X &\in G \end{aligned}$$

Therefore the exponential of any element of $\mathfrak{g}$ is in G. Note that if G is not connected, elements in a connected component different from the identity cannot be expressed as exponentials of a Lie algebra element. If an element in a different connected component could be expressed as e^X for some $X \in \mathfrak{g}$, then e^{tX} would be a path that connects e^X to $\mathbf{1}$, which is a contradiction.

For example, $O(3)$ is just $SO(3)$ where half of the matrices have determinant -1. The determinant -1 matrices are disconnected from the determinant 1 matrices. This is because the determinant map is continuous, and if there were a path from a determinant -1 matrix to a $+1$ matrix, there would then be a continuous function from -1 to 1 that doesn't pass through 0.

The Lie algebra, however, only depends on the group structure in the neighborhood of the identity. Therefore, $\mathfrak{o}(3) = \mathfrak{so}(3)$. We study Lie Algebras because their simple structure is easier to understand than that of the groups. However, as we can see here, Lie algebras can sometimes be blind to the global topological structure of our Lie groups. We always need to be careful.

We are still not done discussing all of the ways that the Lie algebra interacts with the group! It turns out for any $g \in G$ and $X \in \mathfrak{g}$, we have

$$gXg^{-1} \in \mathfrak{g}.$$

This is an interesting and important fact which is easy to prove. For any real t, the matrix

$$e^{tX}$$

is a path in our group that runs through the identity at $t = 0$. Note that

$$ge^{tX}g^{-1}$$

is also a path in our group that runs through the identity. Differentiating by t and evaluating at $t = 0$ will give us the velocity vector of this path at the identity, which is in the tangent space $T_1G = \mathfrak{g}$.

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(ge^{tX}g^{-1}\right)\bigg|_{t=0} &\in \mathfrak{g} \\ = gXe^{tX}g^{-1}\bigg|_{t=0} &\in \mathfrak{g} \\ = gXg^{-1} &\in \mathfrak{g} \end{aligned}$$

As gXg^{-1} is an element of the Lie algebra, then $e^{tY}Xe^{-tY}$, for any $Y \in \mathfrak{g}$, is a path through the Lie algebra. As the velocity vector of a path moving through a vector space is always an element of that vector space itself,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tY}Xe^{-tY}\right)\bigg|_{t=0} &\in \mathfrak{g} \\ = e^{tY}YXe^{-tY} + e^{tY}X(-Y)e^{-tY}\bigg|_{t=0} &\in \mathfrak{g} \\ = YX - XY &\in \mathfrak{g}. \end{aligned}$$

You can also understand the above fact via the conjugation of X by a group element close to the identity.

$$\Big(\mathbf{1} + \epsilon Y + \mathcal{O}(\epsilon^2)\Big)X\Big(\mathbf{1} - \epsilon Y + \mathcal{O}(\epsilon^2)\Big) = X + \epsilon(YX - XY) + \mathcal{O}(\epsilon^2)$$

In either case, we can see that Lie Algebra are not only closed under real linear combinations of the elements: they are also closed under the commutator!

$$\text{For all } X, Y \in \mathfrak{g}, \qquad [X, Y] \in \mathfrak{g}$$

In this context, the commutator is sometimes called the "Lie bracket."

8 The Lie Algebras of $U(n), SL(n, \mathbb{C})$, and $SU(n)$

Let us study the Lie algebra of some simple groups, starting with $U(n)$. For each $U \in U(n)$, $UU^\dagger = \mathbf{1}$. Therefore,

$$\text{for all } X \in \mathfrak{u}(n), \qquad e^{tX} e^{tX^\dagger} = \mathbf{1}.$$

Differentiating this equation gives us a constraint on X:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tX} e^{tX^\dagger}\right)\bigg|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{1}\right)\bigg|_{t=0} \\ \left(e^{tX} X e^{tX^\dagger} + e^{tX} X^\dagger e^{tX^\dagger}\right)\bigg|_{t=0} &= \mathbf{0} \\ X + X^\dagger &= \mathbf{0} \end{aligned}$$

Therefore all elements of $\mathfrak{u}(n)$ must be skew-adjoint, which just means $X = -X^\dagger$.

We now wonder: are *all* skew adjoint matrices in $\mathfrak{u}(n)$? To figure this out, we can use the fact that the dimension of the tangent space of a manifold is the same as the dimension of the manifold itself. (Here, by "dimension" I mean the necessary number of degrees of freedom it takes to parameterize $U(n)$. I don't mean the dimension of the matrices, which is just n.) So what is the dimension of $U(n)$? Well, it lives in the ambient space of complex $n \times n$ matrices, which is $2n^2$ dimensional. The condition $UU^\dagger = \mathbf{1}$ acts as a constraint. If we label the entries of U with n^2 complex numbers, this constraint becomes $\sum_k u_{ik} u^*_{jk} = \delta_{ij}$. When $i = j$ this counts as *one* real constraint: $(a + bi)(a - bi) = 1 \Rightarrow a^2 + b^2 = 1$. There are n ways for i to equal j, so this accounts for n constraints. When $i \neq j$, this counts as *two* real constraints, as both the real part and imaginary part must be 0: $(a + bi)(c + di) = 0 \Rightarrow ac - bd = 0, bc + ad = 0$. There are $n(n-1)/2$ ways for $i \neq j$, so this accounts for $n(n-1)$ real constraints. Our final count for the dimension of $U(n)$ is

$$2n^2 - n - n(n-1) = n^2$$

(By the way, this is a rigorous proof. The constraints cut down on the independent parameters in a small neighborhood of the manifold, and this number is independent of the small neighborhood.)

Let us now count the dimensions of the space of skew-adjoint matrices. Skew-adjoint matrices are specified by the entries in the upper-triangular portion of the matrix. Furthermore, the diagonal entries must be pure imaginary. We have n real degrees of freedom for the diagonal and $2 \times (1 + \ldots + (n-1)) = n(n-1)$ for the rest of the upper triangular portion. Therefore, the dimension of the space of skew-adjoint matrices is

$$n + n(n-1) = n^2.$$

The dimension of both is n^2! Now, the dimension of the tangent space of any manifold is always equal to the dimension of the manifold itself. Therefore, the tangent space cannot be a subspace of the vector space of skew-adjoint matrices, and must be the whole space.

$$\mathfrak{u}(n) = \{X \in M(n, \mathbb{C}) \text{ such that } X^\dagger = -X\}$$

(By the way, $M(n, \mathbb{C})$ is the vector space of $n \times n$ complex matrices. It is not a group under matrix multiplication because it contains non-invertible matrices with determinant 0.)

A note: often in physics we prefer to work with self-adjoint matrices instead of skew-adjoint ones. Luckily, there is a bijection between self-adjoint matrices and skew-adjoint given by multiplication by i.

$$X = -iA$$
$$X^\dagger = -X \to A^\dagger = A$$
$$e^X = e^{-iA}$$

Let's move onto the Lie algebra of $SL(n, \mathbb{C})$, the "special linear group," the group of complex $n \times n$ matrices with determinant 1. For an element $X \in \mathfrak{sl}(n)$, we have

$$e^X \in SL(n, \mathbb{C}).$$

From the matrix identity

$$g = e^X \quad \implies \quad \det(g) = e^{\operatorname{tr} X}$$

we can see that the elements of $\mathfrak{sl}(n)$ must necessarily be traceless. We will count dimensions again to show that this is also a sufficient condition for a matrix to be an element of $\mathfrak{sl}(n)$.

For a complex matrix to have a determinant of 1 is to impose 2 real constraints on it, as the determinant is just a function of all the matrix elements. These two constraints are $\text{Re}(\det) = 1$ and $\text{Im}(\det) = 0$. Therefore the dimension of $SL(n, \mathbb{C})$ is

$$2n^2 - 2.$$

Likewise, for a complex matrix to be traceless is to impose 2 real constraints on it. Namely $\text{Re}(\text{tr}) = 0$ and $\text{Im}(\text{tr}) = 0$. Therefore the dimension of the space of trace 0 matrices is

$$2n^2 - 2.$$

The dimensions match! Therefore

$$\mathfrak{sl}(n) = \{X \in M(n, \mathbb{C}) \text{ such that } \operatorname{tr} X = 0\}.$$

Let us now turn to $SU(n)$, the group of unitary matrices with determinant 1. This group is clearly just the intersection of $U(n)$ and $SL(n, \mathbb{C})$. Therefore the tangent space at the identity is just the intersection of the tangent spaces of both $U(n)$ and $SL(n,\mathbb{C})$. We can immediately see that the Lie algebra $\mathfrak{su}(n)$ is just the space of traceless skew-adjoint matrices.

$$\mathfrak{su}(n) = \{X \in M(n,\mathbb{C}) \text{ such that } X^\dagger = -X \text{ and } \operatorname{tr} X = 0\}$$

9 Lie Algebra Representations

If we have a Lie group homomorphism

$$\pi : G_1 \to G_2$$

then, for some $X \in \mathfrak{g}_1$, $\pi(e^{tX})$ is a path in G_2 that passes through the identity at $t = 0$. The velocity vector of this path at $t = 0$ will therefore be an element of $\mathfrak{g}_2$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\, \pi(e^{tX})\bigg|_{t=0} \in \mathfrak{g}_2$$

Therefore, we can define a map

$$\pi' : \mathfrak{g}_1 \to \mathfrak{g}_2$$

by

$$\pi'(X) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\, \pi(e^{tX})\bigg|_{t=0}.$$

Based on the definition above, we have the following identity:

$$\pi(e^{tX}) = e^{t\pi'(X)}.$$

The proof of this is simple. The homomorphism property of π allows us to create a differential equation that $\pi(e^{tX})$ satisfies. The "initial condition" $\pi(\mathbf{1}) = \mathbf{1}$ allows us to solve it uniquely:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\pi(e^{tX}) &= \frac{\mathrm{d}}{\mathrm{d}s}\, \pi(e^{(t+s)X})\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\, \pi(e^{sX})\bigg|_{s=0} \pi(e^{tX}) \\
&= \pi'(X)\pi(e^{tX}) \\
\implies \pi(e^{tX}) &= e^{t\pi'(X)}
\end{aligned}$$

The equation $\pi(e^{tX}) = e^{t\pi'(X)}$ is what we will actually use when we want to find π' based on π.

We can also show that $\pi'(gXg^{-1}) = \pi(g)\pi'(X)\pi(g)^{-1}$:

$$\begin{aligned}
\pi'(gXg^{-1}) &= \frac{\mathrm{d}}{\mathrm{d}t}\, \pi(e^{tgXg^{-1}})\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\, \pi(g)\pi(e^{tX})\pi(g)^{-1}\bigg|_{t=0} \\
&= \pi(g)\pi'(X)\pi(g)^{-1}
\end{aligned}$$

Using these two equations, we can show that π' satisfies two important properties. The first is linearity:

$$\begin{aligned}
\pi'(\alpha X + \beta Y) &= \frac{\mathrm{d}}{\mathrm{d}t}\, \pi(e^{t(\alpha X+\beta Y)})\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\, \pi(e^{t\alpha X})\pi(e^{t\beta Y})\pi(e^{\mathcal{O}(t^2)})\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\, e^{t\alpha\pi'(X)}e^{t\beta\pi'(Y)}\pi(e^{\mathcal{O}(t^2)})\Big|_{t=0} \\
&= \alpha\pi'(X) + \beta\pi'(Y)
\end{aligned}$$

The second is that π' preserves the Lie brackets:

$$\begin{aligned}
\pi'([X,Y]) &= \pi'\Big(\frac{\mathrm{d}}{\mathrm{d}t}\, (e^{tX}Ye^{-tX})\Big|_{t=0}\Big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\, \pi'(e^{tX}Ye^{-tX})\Big|_{t=0} \qquad \text{(by linearity)} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\, \pi(e^{tX})\pi'(Y)\pi(e^{-tX})\Big|_{t=0} \\
&= [\pi'(X), \pi'(Y)]
\end{aligned}$$

These two properties are very important. Instead of being derived from a given π', they are usually a starting point for definitions:

Any map $\pi' : \mathfrak{g}_1 \to \mathfrak{g}_2$ is said to be a **Lie algebra homomorphism** if

1. $\pi'(\alpha X + \beta Y) = \alpha\pi'(X) + \beta\pi'(Y)$ for all $\alpha, \beta \in \mathbb{R}, X, Y \in \mathfrak{g}_1$.
2. $\pi'([X,Y]) = [\pi'(X), \pi'(Y)]$ for all $X, Y \in \mathfrak{g}_1$.

Such a homomorphism π' is a **Lie Algebra representation** when $\mathfrak{g}_2$ has a linear action on a vector space. Because we are only considering Lie groups that are matrix groups, all of our Lie algebra elements are matrices. Any such Lie algebra homomorphisms will automatically be a Lie algebra representation. One last definition: A Lie algebra representation is called "unitary" if

$$\pi'(X)^\dagger = -\pi'(X) \text{ for all } X \in \mathfrak{g}_1.$$

10 Classifying the Irreducible Representations of $U(1)$

Let us leave this Lie algebra talk for a moment to address an important question. What are the irreducible representations of $U(1)$?

$U(1)$ is the simplest Lie group. It's also known as the "circle group," because it's all the unitary 1×1 matrices, i.e. matrices of the form

$$[e^{i\theta}]$$

i.e., a circle. This group is abelian. Therefore, from Corollary 1 of Schur's Lemma, we know that all of the irreducible representations of $U(1)$ must be 1-dimensional.

But 1×1 matrices are just complex numbers. Therefore we should have no problem differentiating an irreducible representation $\pi : U(1) \to GL(1, \mathbb{C})$.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\theta}\pi(e^{i\theta}) &= \lim_{\Delta\theta \to 0} \frac{\pi(e^{i(\theta+\Delta\theta)}) - \pi(e^{i\theta})}{\Delta\theta} \\
&= \pi(e^{i\theta}) \lim_{\Delta\theta \to 0} \frac{\pi(e^{i\Delta\theta}) - \pi(1)}{\Delta\theta} \\
&= \pi(e^{i\theta}) \frac{\mathrm{d}}{\mathrm{d}\theta} \pi(e^{i\theta})\Big|_{\theta=0} \\
&= R\pi(e^{i\theta}) \\
\implies \pi(e^{i\theta}) &= e^{iR\theta}
\end{aligned}$$

$R = \frac{\mathrm{d}}{\mathrm{d}\theta} \pi(e^{i\theta})\big|_{\theta=0}$ is a constant that depends uniquely on the irreducible representation in question. The homomorphism property requires that $e^{iR2\pi} = 1$. This implies that R must be an integer.

We have now completely classified the irreducible representations of $U(1)$. Irreps are labeled by integers. For every $k \in \mathbb{Z}$, we have an irrep given by

$$\pi_k(e^{i\theta}) = e^{ik\theta}.$$

Therefore, for any n-dimensional representation of $U(1)$, there is a basis where the representation is diagonalized:

$$\pi(e^{i\theta}) = \begin{pmatrix} e^{ik_1\theta} & 0 & \cdots & 0 \\ 0 & e^{ik_2\theta} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{ik_n\theta} \end{pmatrix} = e^{iK\theta}$$

where $K = \mathrm{diag}(k_1, k_2, \dots, k_n)$. Note that K commutes with all elements of the representation. K is often called the "charge operator" for reasons that should currently be mysterious. For now, just appreciate that the integers lie at the heart of $U(1)$ representations.

11 Lie Algebra Complexifications

"Complexification" is a confusing thing. Let me explain what it is.

Only real linear combinations of Lie algebra elements remain in the Lie algebra. This is definitional. We defined Lie algebras as the tangent spaces of of Lie groups at the identity. Tangent spaces of manifolds are real vector spaces. Therefore, Lie algebras are real vector spaces too.

Lets look at a specific example. We already know that $\mathfrak{su}(2)$ is the vector space of 2×2 traceless skew-adjoint matrices.

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} -ix & -y - iz \\ y - iz & ix \end{pmatrix} \text{ for all } x, y, z \in \mathbb{R} \right\}$$

Therefore, $\mathfrak{su}(2)$ is the span of three matrices over $\mathbb{R}$:

$$\mathfrak{su}(2) = \mathrm{span}_{\mathbb{R}}\{X_1, X_2, X_3\}$$

where

$$X_1 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -\frac{i}{2}\sigma_1 \qquad X_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -\frac{i}{2}\sigma_2 \qquad X_3 = \frac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = -\frac{i}{2}\sigma_3.$$

The matrices X_j are indeed complex matrices. However, only real linear combinations of the X_j are in $\mathfrak{su}(2)$.

It's not hard to see why we cannot allow complex linear combinations. Take, for example, the matrix iX_1. This matrix is self-adjoint, not skew adjoint. Therefore iX_1 cannot be an element of $\mathfrak{su}(2)$.

Even though complex linear combinations of X_j are not in $\mathfrak{su}(2)$, we can define the *complexified Lie algebra* $\mathfrak{su}(2)_{\mathbb{C}}$

$$\mathfrak{su}(2)_{\mathbb{C}} \equiv \text{span}_{\mathbb{C}}\{X_1, X_2, X_3\}$$

such that complex linear combinations are allowed. We can now see that $\mathfrak{su}(2)_{\mathbb{C}}$ is 6-dimensional instead of 3-dimensional.

$$\mathfrak{su}(2)_{\mathbb{C}} = \left\{ \begin{pmatrix} -u & v \\ w & u \end{pmatrix} \text{ for all } u, v, w \in \mathbb{C} \right\}$$

Defining the self adjoint matrices $S_j \in \mathfrak{su}(2)_{\mathbb{C}}$ by

$$S_j \equiv iX_j$$

we can express all elements of $\mathfrak{su}(2)_{\mathbb{C}}$ as real linear combinations of six matrices.

$$\mathfrak{su}(2)_{\mathbb{C}} = \text{span}_{\mathbb{R}}\{X_1, X_2, X_3, S_1, S_2, S_3\}$$

Elements of $\mathfrak{su}(2)_{\mathbb{C}}$ will no longer exponentiate to elements in $SU(2)$. However, from the statement above, we can see that $\mathfrak{su}(2)_{\mathbb{C}}$ is just the real vector space of traceless 2×2 complex matrices. Therefore we can see that

$$\mathfrak{su}(2)_{\mathbb{C}} = \mathfrak{sl}(2).$$

Did that clear things up? In the next part we will use the complexification of $\mathfrak{su}(2)$ in order to classify all irreducible representations of $SU(2)$.

I should mention that the proscription I presented here of just *allowing* complex linear combinations to make a complexified Lie algebra is not general. That would imply, for example, that the complexification of $\mathfrak{sl}(2)$ is just $\mathfrak{sl}(2)$ again. This is not the case. Complexification will always double the dimension of the Lie algebra. In general one must do something more sophisticated to complexify a Lie algebra. For Lie algebras where $\mathfrak{g} \cap i\mathfrak{g} = 0$, however, this isn't necessary.

Note that when we "complexify" our Lie algebra in the simple way, i.e. allowing complex combinations of the basis elements, a Lie algebra representation π' can also be extended in the obvious way by allowing complex numbers to filter through π'. That is, we define $\pi'(iX) = i\pi'(X)$. As the Lie groups we consider all act on complex vector spaces, this is well-defined.

12 Classifying the Irreducible Representations of $SU(2)$

Let us now rigorously classify all of the irreducible representations of $SU(2)$. So far we have only done this for the much simpler group $U(1)$. We will rely heavily on what we know about $U(1)$, Lie algebras representations, and Lie algebra complexifications.

In the last part we showed that $\mathfrak{su}(2)$ is spanned by the three matrices X_j. They are called the "generators" of $SU(2)$. They satisfy the following commutation relations:

$$[X_1, X_2] = X_3 \qquad [X_2, X_3] = X_1 \qquad [X_3, X_1] = X_2.$$

The complexified Lie algebra $\mathfrak{su}(2)_{\mathbb{C}}$ introduces the self adjoint matrices $S_j = iX_j$ into the mix. As $X_j = -\frac{i}{2}\sigma_j$, $S_j = \frac{1}{2}\sigma_j$. The matrices S_j satisfy the following commutation relations:

$$[S_1, S_2] = iS_3 \qquad [S_2, S_3] = iS_1 \qquad [S_3, S_1] = iS_2.$$

Let us now define the "raising and lowering operators" $S_+, S_- \in \mathfrak{su}(2)_{\mathbb{C}}$:

$$S_+ \equiv S_1 + iS_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad S_- \equiv S_1 - iS_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The important thing about these operators is how they interact with S_3.

$$\begin{aligned} [S_3, S_+] &= [S_3, S_1] + i[S_3, S_2] \\ &= iS_2 + i(-iS_1) \\ &= S_+ \end{aligned}$$

$$\begin{aligned} [S_3, S_-] &= [S_3, S_1] - i[S_3, S_2] \\ &= iS_2 - i(-iS_1) \\ &= -S_- \end{aligned}$$

$$\begin{aligned} [S_+, S_-] &= [S_1, S_1] + i[S_2, S_1] - i[S_1, S_2] - i^2[S_2, S_2] \\ &= i(-iS_3) - i(iS_3) \\ &= 2S_3 \end{aligned}$$

As $iS_3 \in \mathfrak{su}(2)$, $e^{i2\theta S_3} \in SU(2)$. More specifically,

$$e^{i\theta 2S_3} = e^{i\theta\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

We can see that the subgroup of $SU(2)$ "generated" by iS_3 is isomorphic to $U(1)$.

Therefore, if we have a representation π of $SU(2)$

$$\pi : SU(2) \to GL(n, \mathbb{C})$$

we can restrict π to this $U(1)$ subgroup and turn π into a representation of $U(1)$. Using our knowledge of representations of $U(1)$, we know that in some basis $\pi(e^{i\theta 2S_3})$ can always be written as follows:

$$\pi(e^{i\theta 2S_3}) = \begin{pmatrix} e^{ik_1\theta} & 0 & \cdots & 0 \\ 0 & e^{ik_2\theta} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{ik_n\theta} \end{pmatrix}.$$

Differentiating by θ, we can see where the associated Lie algebra representation π' sends S_3:

$$\pi'(S_3) = \begin{pmatrix} \frac{k_1}{2} & 0 & \cdots & 0 \\ 0 & \frac{k_2}{2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{k_n}{2} \end{pmatrix}.$$

The half-integers $\frac{k_1}{2}, \ldots, \frac{k_n}{2}$ depend explicitly on π, the representation of $SU(2)$ in question. The integers k are called the "weights" of π. If π is an n-dimensional irreducible representation then these weights obey a specific pattern which we will derive shortly.

First, note the following interesting property: If v_k is an eigenvector of $\pi'(S_3)$ with eigenvalue $\frac{k}{2}$, then $\pi'(S_+)v_k$ is an eigenvector of $\pi'(S_3)$ with eigenvalue $(\frac{k}{2}+1)$ and $\pi'(S_-)v_k$ is an eigenvector of $\pi'(S_3)$ with eigenvalue $(\frac{k}{2}-1)$.

$$\pi'(S_3)v_k = \tfrac{k}{2}v_k \quad \implies \quad \begin{aligned} \pi'(S_3)\pi'(S_+)v_k &= (\tfrac{k}{2}+1)\pi'(S_+)v_k \\ \pi'(S_3)\pi'(S_-)v_k &= (\tfrac{k}{2}-1)\pi'(S_-)v_k \end{aligned}$$

This can be derived straight from the commutation relations.

$$\begin{aligned} \pi'(S_3)\pi'(S_+)v_k &= \Big(\pi'(S_+)\pi'(S_3) + [\pi'(S_3), \pi'(S_+)]\Big)v_k \\ &= \Big(\pi'(S_+)\pi'(S_3) + \pi'([S_3, S_+])\Big)v_k \\ &= \Big(\pi'(S_+)\pi'(S_3) + \pi'(S_+)\Big)v_k \\ &= \pi'(S_+)\big(\pi'(S_3) + 1\big)v_k \\ &= (\tfrac{k}{2}+1)\pi'(S_+)v_k \end{aligned}$$

$$\begin{aligned} \pi'(S_3)\pi'(S_-)v_k &= \Big(\pi'(S_-)\pi'(S_3) + \pi'([S_3, S_-])\Big)v_k \\ &= \pi'(S_-)\big(\pi'(S_3) - 1\big)v_k \\ &= (\tfrac{k}{2}-1)\pi'(S_-)v_k \end{aligned}$$

Note that for any finite dimensional representation π there will be a "highest weight vector" v_{k_H}. This vector is an eigenvector of $\pi'(S_3)$ with the largest eigenvalue $k_H/2$. As there is no vector with a larger eigenvalue, we must have

$$\pi'(S_+)v_{k_H} = 0.$$

There will also be a lowest weight vector v_{k_L} that satisfies

$$\pi'(S_-)v_{k_L} = 0.$$

Say that $\pi : SU(2) \to GL(V)$ is a finite dimensional irreducible representation of $SU(2)$ where V is a complex vector space. There will be a highest weight vector $v_{k_H} \in V$. I claim that the following subspace $W \subset V$ is closed under the action of all Lie algebra representation elements $\pi'(X)$:

$$W \equiv \mathrm{span}_{\mathbb{C}}\Big\{v_{k_H}, \pi'(S_-)v_{k_H}, \pi'(S_-)^2 v_{k_H}, \pi'(S_-)^3 v_{k_H}, \dots\Big\}$$

(By "closed" I mean that for all $w \in W$ and $X \in \mathfrak{su}(2)$, we have $\pi'(X)w \in W$.)

Proof: As W is a complex vector space, all vectors $w \in W$ and Lie algebra representation elements $\pi'(X)$ satisfy $\pi'(X)w \in W$ if and only if $i\pi'(X)w \in W$. Therefore, if we show that W is closed under the action of the *complexified* Lie algebra $\mathfrak{su}(2)_{\mathbb{C}}$, we are done.

A basis of $\mathfrak{su}(2)_{\mathbb{C}}$ is given by S_+, S_-, and S_3:

$$\mathfrak{su}(2)_{\mathbb{C}} = \mathrm{span}_{\mathbb{C}}\Big\{S_+, S_-, S_3\Big\}$$

We can clearly see that W is closed under the action of $\pi'(S_-)$ and $\pi'(S_3)$. Therefore all we must do is check that it is closed under the action of $\pi'(S_+)$. This can be achieved from the commutation relations. As

$$[\pi'(S_+), \pi'(S_-)] = 2\pi'(S_3)$$

we have

$$\begin{aligned}
\pi'(S_+)\pi'(S_-)^m v_{k_H} &= \big(\pi'(S_-)\pi'(S_+) + [\pi'(S_+), \pi'(S_-)]\big)\pi'(S_-)^{m-1} v_{k_H} \\
&= \big(\pi'(S_-)\pi'(S_+) + 2\pi'(S_3)\big)\pi'(S_-)^{m-1} v_{k_H} \\
&= \sum_{j=0}^{m-1} \pi'(S_-)^{m-1-j} 2\pi'(S_3)\pi'(S_-)^j v_{k_H} \\
&= \Big(\sum_{j=0}^{m-1} k_H - 2j\Big)\pi'(S_-)^{m-1} v_{k_H} \\
&\subset W. \quad \square
\end{aligned}$$

A schematic of how $\pi'(S_+)$, $\pi'(S_-)$ and $\pi'(S_3)$ act on these basis vectors is depicted in Figure 4.

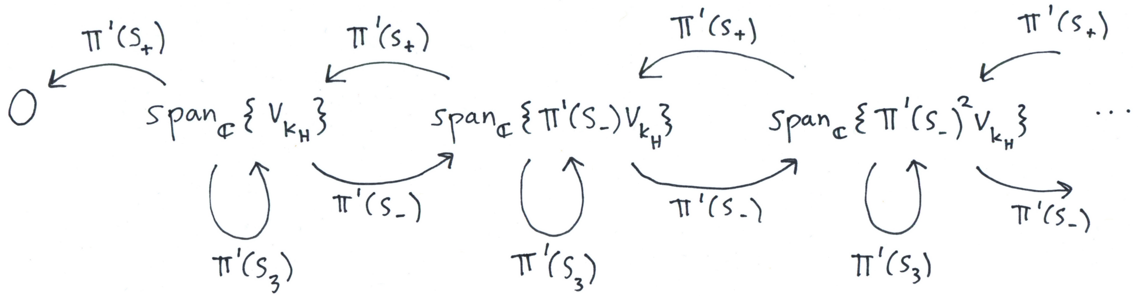

Figure 4: The subspaces (called "weight spaces") of W and how different elements act on them.

It can be explicitly checked that every matrix in $SU(2)$ can be expressed as e^X for some $X \in \mathfrak{su}(2)$ via matrix diagonalization. As $\pi(e^X) = e^{\pi'(X)}$, it can be seen that any Lie group representation element $\pi(g)$ is a sum of products of Lie algebra representation elements $\pi'(X)$ (and $\mathbf{1}$).

$$\pi(g) = \mathbf{1} + \pi'(X) + \frac{1}{2}\pi'(X)^2 + \frac{1}{6}\pi'(X)^3 + \dots$$

Therefore, we can see that W is also closed under the action of $\pi(g)$ for any $g \in SU(2)$. Therefore W is a sub-representation of V. As V is irreducible, $W = V$.

We have therefore proved that *every* finite-dimensional irreducible representation of $SU(2)$ is spanned by repeatedly lowering a highest weight vector. Therefore, we know that our weights will be, with multiplicity 1,

$$k_H, k_H - 2, k_H - 4, \dots, k_L + 4, k_L + 2, k_L$$

where $k_H - k_L$ is an even number.

If $V = \mathbb{C}^n$ there must be n distinct weights (as each weight has multiplicity 1). We can see that if $n = 1$ then the only possible irreducible representation is the trivial one $\pi(g) = 1$. This is because V is spanned by exactly one vector, say v, which must be both a highest weight vector and a lowest weight vector. That is, $\pi'(S_+)v = 0$ and $\pi'(S_-)v = 0$. So $\pi'(S_+) = \pi'(S_-) = 0$. Furthermore, $\pi'(S_3) = \frac{1}{2}[\pi'(S_+), \pi'(S_-)] = \frac{1}{2}[0, 0] = 0$. As S_+, S_-, and S_3 span the Lie algebra, $\pi(e^X) = e^{\pi'(X)} = 1$.

This allows us to prove that for any irrep $\pi : SU(2) \to GL(\mathbb{C}^n)$,

$$\det \pi(g) = 1 \text{ for all } g \in SU(2).$$

This is because the determinant map $\det : GL(V) \to \mathbb{C} = GL(\mathbb{C})$ is itself a representation. Therefore, if we compose it with π, we receive a 1-dimension representation of $SU(2)$.

$$(\det \circ \pi) : SU(2) \xrightarrow{\pi} GL(\mathbb{C}^n) \xrightarrow{\det} GL(\mathbb{C})$$

As all 1-dimensional irreps of $SU(2)$ are trivial, we have shown $\det \pi(g) = 1$.

For the element $e^{i\theta 2S_3} \in SU(2)$ and an irrep $\pi : SU(2) \to GL(\mathbb{C}^n)$, we have

$$\pi(e^{i\theta 2S_3}) = \mathrm{diag}(e^{ik_H\theta}, e^{i(k_H-2)\theta}, \dots, e^{i(k_L+2)\theta}, e^{ik_L\theta}).$$

As $\det \pi(e^{i\theta 2S_3}) = 1$, we can see that the sum of all weights must be 0. Therefore, the weights must be

$$n, n-2, n-4, \dots, 4-n, 2-n, -n.$$

We have now shown what the weights must be for an n-dimensional irrep of $SU(2)$. Furthermore, we have explicitly expressed $\pi'(S_+)$, $\pi'(S_-)$, and $\pi'(S_3)$ in terms of the basis vectors $\pi'(S_-)^m v_H$. Therefore our Lie algebra representation π' is uniquely determined by the dimension of the irrep. As a representation of $SU(2)$ is specified by its Lie algebra representation, we have shown that there is exactly one irrep of $SU(2)$ for any dimension.

We usually refer to these irreducible representations as the "spin j" representation, where j is a half integer and the dimension of the representation is $2j + 1$. Note that j is the highest weight of the representation.

13 The Spin 1 Representation of $SU(2)$

The spin 0 representation of $SU(2)$ is just the trivial representation $\pi(g) = 1$, and the spin $\frac{1}{2}$ representation is just the defining representation $\pi(g) = g$. The first non-trivial irrep of $SU(2)$ is therefore the spin 1 representation, which maps elements of $SU(2)$ into $GL(\mathbb{C}^3)$. Using our discussion in the last part, we will see what exactly it is.

The highest weight is $k_H = 2$. We will use basis $\{\sqrt{2}v_2, \pi'(S_-)v_2, \frac{1}{\sqrt{2}}\pi'(S_-)^2 v_2\}$. Note the $\sqrt{2}$'s, which were chosen for convenience. This basis is often called the "Zeeman basis." We know that

$$\pi'(S_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and

$$\pi'(S_-) = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \qquad \pi'(S_+) = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}$$

where we used the expression $\pi'(S_+)\pi'(S_-)^m v_{k_H} = \left(\sum_{j=0}^{m-1} k_H - 2j\right)\pi'(S_-)^{m-1} v_{k_H}$ which implies that $\pi'(S_+)\pi'(S_-)v_2 = 2v_2$ and $\pi'(S_+)\pi'(S_-)^2 v_2 = 2\pi'(S_-)v_2$. Using the fact that

$$S_1 = \frac{1}{2}(S_+ + S_-) \qquad S_2 = -\frac{i}{2}(S_+ - S_-)$$

we can write π' for the 1 representation as follows:

$$\pi'(S_1) = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \qquad \pi'(S_2) = \begin{pmatrix} 0 & -\frac{i}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ 0 & \frac{i}{\sqrt{2}} & 0 \end{pmatrix} \qquad \pi'(S_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Every element of the group representation can then be expressed as

$$\pi(e^{-i\vec{\theta}\cdot\vec{S}}) = e^{-i\vec{\theta}\cdot\pi'(\vec{S})} = e^{\vec{\theta}\cdot\pi'(\vec{X})}$$

for some $\vec{\theta}$, remembering that $X_j = -iS_j$ is the true basis of $\mathfrak{su}(2)$. Because we cleverly used the Zeeman basis, this Lie algebra representation is unitary. Another basis would have produced an equivalent representation that isn't necessarily unitary.

14 An Explicit Construction of Unitary $SU(2)$ Representations

These raising and lowering operators are great and all, but if you want to explicitly construct unitary representations of $SU(2)$ there's a pretty easy way to do it.

Consider the space of homogenous complex polynomials in 2 variables of degree m.

$$P(z_1, z_2)$$

"Homogenous" just means that for some $m \in \mathbb{Z}$

$$P(\lambda z_1, \lambda z_2) = \lambda^m P(z_1, z_2) \text{ for all } \lambda \in \mathbb{C}.$$

This is clearly a complex vector space. For example, the space of homogenous complex polynomials of degree 2 is spanned by the following three polynomials:

$$\frac{z_1^2}{\sqrt{2}} \qquad z_1 z_2 \qquad \frac{z_2^2}{\sqrt{2}}.$$

A hermitian inner product on this space is given by

$$\langle f, g\rangle \equiv \frac{1}{\pi^2}\int_{\mathbb{C}^2} \mathrm{d}^2 z_1 \mathrm{d}^2 z_2 e^{-|z_1|^2-|z_2|^2} f^*(z_1, z_2) g(z_1, z_2).$$

It can be checked that, with respect to this inner product, an orthonormal basis is given by monic polynomials of the form

$$\frac{z_1^j z_2^k}{\sqrt{j!k!}}$$

where $j + k = m$. This is clearly an $m + 1$ dimensional space.

This vector space of polynomials has a natural action under $SU(2)$ if we imagine that our polynomials take in column vectors from $\mathbb{C}^2$:

$$\big(\pi(g)f\big)\left(\begin{smallmatrix} z_1 \\ z_2 \end{smallmatrix}\right) = f\left(g^{-1}\left(\begin{smallmatrix} z_1 \\ z_2 \end{smallmatrix}\right)\right)$$

A change of variables in the inner product shows that this group action is unitary. The highest weight vector is the polynomial

$$\frac{z_1^m}{\sqrt{m!}}$$

and the raising and lowering operators are

$$\pi'(S_+) = z_1 \frac{\partial}{\partial z_2} \qquad\qquad \pi'(S_-) = z_2 \frac{\partial}{\partial z_1}.$$

All of the weight vectors are given by the monic polynomials:

$$\pi'(S_3)\frac{z_1^{m-k} z_2^k}{\sqrt{(m-k)!k!}} = \frac{1}{2}(m-2k)\frac{z_1^{m-k} z_2^k}{\sqrt{(m-k)!k!}}$$

This is done in more detail in Woit's book.

15 The Odd Couple: $SU(2)$ and $SO(3)$

$SO(3)$, the group of 3×3 special orthogonal matrices, has a very weird and intimate connection with $SU(2)$. Let's get to the bottom of it.

Any matrix in $SO(3)$ can be expressed as the composition of a rotation around the x, y, and z axes.

$$R(\theta, \varphi, \psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

It's not hard to see that the following three matrices form a basis of $\mathfrak{so}(3)$.

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad L_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad L_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

They satisfy the following commutation relations.

$$[L_1, L_2] = L_3 \qquad [L_2, L_3] = L_1 \qquad [L_3, L_1] = L_2$$

This looks a lot like the commutation relations for $\mathfrak{su}(2)$, spanned by $X_j = -\frac{i}{2}\sigma_j$.

$$[X_1, X_2] = X_3 \qquad [X_2, X_3] = X_1 \qquad [X_3, X_1] = X_2$$

This allows us to create a Lie algebra homomorphism

$$\pi' : \mathfrak{su}(2) \to \mathfrak{so}(3)$$

defined by

$$\pi'(X_j) = L_j.$$

We can exponentiate this Lie algebra homomorphism into a Lie group homomorphism

$$\pi : SU(2) \to SO(3)$$

defined by

$$\pi(e^{\vec{\theta}\cdot\vec{X}}) = e^{\vec{\theta}\cdot\pi'(\vec{X})} = e^{\vec{\theta}\cdot\vec{L}}.$$

Of course, we would need to check that this truly is a group homomorphism and show that $\pi(g_1 g_2) = \pi(g_1)\pi(g_2)$. This could follow from a slightly sketchy application of the Baker-Campbell-Hausdorff formula, using the fact that π' is a Lie algebra homomorphism. There is, however, a more useful way to express the $SU(2) \to SO(3)$ homomorphism.

Here's how it works. First, express a vector $\vec{x}$ as a traceless self-adjoint 2×2 matrix. The space of such matrices is three dimensional, so this is an isomorphism.

$$(x, y, z) \longleftrightarrow x\sigma_x + y\sigma_y + z\sigma_z = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}$$

The action of $\pi(g)$ on $\vec{x}$ is then given by cojugation of this matrix by g. As the conjugation of a self-adjoint matrix by a unitary one is still self-adjoint, the resulting matrix can again be uniquely identified with vector $\vec{x}$. Let's do an example.

$$
\begin{aligned}
\pi\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta}\end{smallmatrix}\right)\vec{x} &\leftrightarrow \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta}\end{pmatrix}\begin{pmatrix} z & x-iy \\ x+iy & -z\end{pmatrix}\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta}\end{pmatrix}^{-1} \\
&= \begin{pmatrix} e^{i\theta}z & e^{i\theta}(x-iy) \\ e^{-i\theta}(x+iy) & -e^{-i\theta}z\end{pmatrix}\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta}\end{pmatrix} \\
&= \begin{pmatrix} z & e^{i2\theta}(x-iy) \\ e^{-i2\theta}(x+iy) & -z\end{pmatrix} \\
&= \begin{pmatrix} z & (xc2\theta+ys2\theta)-i(-xs2\theta+yc2\theta) \\ (xc2\theta+ys2\theta)+i(-xs2\theta+yc2\theta) & -z\end{pmatrix} \\
&\leftrightarrow \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix} x \\ y \\ z\end{pmatrix}
\end{aligned}
$$

In this way we can explicitly find which $SO(3)$ matrix any $SU(2)$ matrix maps to. (We also can easily notice an interesting fact: both g and $-g$ will map to the same element of our homomorphism. That is, $\pi(-g) = \pi(g)$.)

But remember, we have already classified all representations of $SU(2)$! Therefore, $SO(3)$ must necessarily be the spin 1 representation of $SU(2)$. There's no other option. It's just a change of basis away from the spin 1 representation we've already investigated.

Here is where the fun begins. Let us try to *picture* $SO(3)$ as a space in its own right, warts and all. Every rotation is specified by an angle of rotation $\theta \in [0, \pi]$ and a unit vector $\hat{n}$ that serves as the axis of rotation. We can therefore take the vector $\theta\hat{n}$ to specify a rotation. This is depicted in Figure 5.

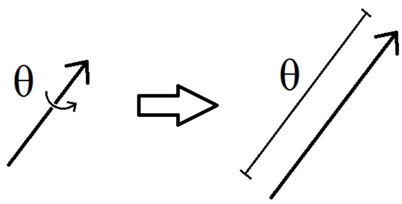

Figure 5: Every rotation is specified by a unit vector an an angle θ. Alternatively, every rotation is specified by a vector with magnitude θ.

The vectors of magnitude π, however, are not unique. After all, a rotation by π around $\hat{n}$ is the same a rotation by π around $-\hat{n}$! That is, $\pi\hat{n}$ and $-\pi\hat{n}$ correspond to the same rotation.

All of these vectors are contained within the sphere of radius π. As we have seen, in order to ensure that every point in this sphere corresponds to exactly one rotation, we must regard every point on the boundary of this sphere to be the same as its antipodal point!

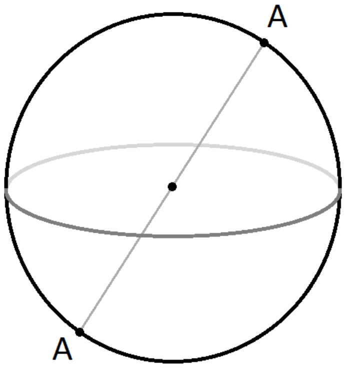

Figure 6: Every element of $SO(3)$ can be identified uniquely by a point in a sphere of radius π as long as we regard antipodal points on the boundary as the same point. Here, the two points labeled "A" are actually the same point.

For those of you well versed in topology, $SO(3)$ is the 3-disk glued to $\mathbb{R}P^2$ in a particular way. You may recognize this space as $\mathbb{R}P^3$! Apparently, $SO(3)$ and $\mathbb{R}P^3$ are one in the same!

What is the fundamental group of $SO(3)$? An application of van Kampen's theorem (or a quick Google search) reveals $\pi_1\big(SO(3)\big) = \pi_1(\mathbb{R}P^3) = \mathbb{Z}_2$. Apparently, up to continuous deformation, there is only one non-contractible path in $SO(3)$! Can we picture it? You bet!

The center of our sphere corresponds to the identity element of $SO(3)$, as it is the rotation that rotates by an angle of 0. Start at the center of the sphere and start moving in any direction. Eventually you will hit the boundary and pop out the other side. Keep moving in your original direction. When you get back to the center, stop. This is a non-contractible loop. Up to continuous deformation, it is the *only* non-contractible loop. It is depicted in Figure 7. Because there is a non-contractible loop, we say that $SO(3)$ is not "simply connected."

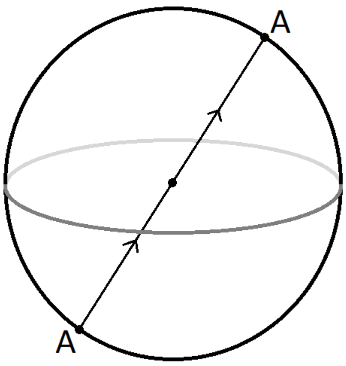

Figure 7: The non-contractible loop in $SO(3)$. This path starts from the origin, then proceeds to the boundary. It then comes out the other side and returns to the starting position.

It seems intuitive that this loop is non-contractible. We can't move the two "manifestations" of A closer to each other, which is what we would have to do to contract the loop.

What happens if we go around this loop twice? That is, what if, when we return to the origin, we go to the boundary again and come out the other side again before returning to the origin? This is loop is depicted in Figure 8, deformed slightly so it is easier to understand. Instead of re-entering the boundary at A again, it enters at an adjacent point B.

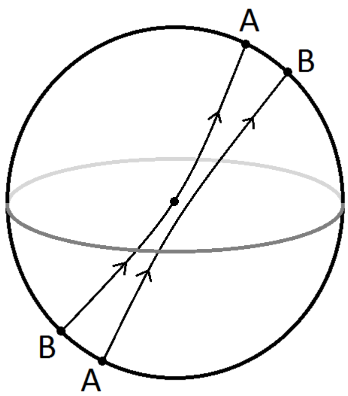

Figure 8: This is the loop obtained by going around the non-contractible loop twice. It starts from the origin, proceeds to A, comes out the other side and proceeds to an adjacent point B before returning to the origin.

The interesting thing is that this loop is actually contractible! We can continuously deform it until it is the "trivial loop," the loop that starts and the origin and doesn't move. The trick is that you have to play with B. While it starts out close to A, you have to move it around the boundary so it leaves A and the returns. This is depicted in Figure 9.

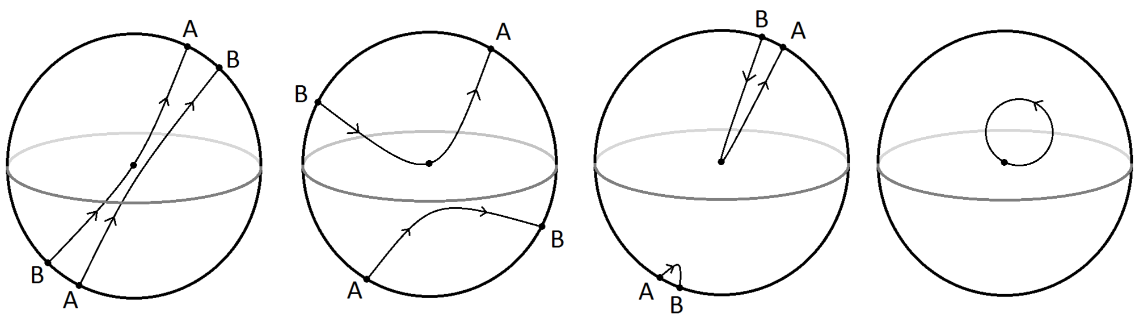

Figure 9: How to contract the loop in Figure 8. The loop in the final diagram can easily be contracted into the "trivial" loop.

Now, one topological space X is said to cover another space Y if there is a continuous map $p : X \to Y$ that is surjective.

Algebraic topology tells us something interesting. It tells us that there is a unique simply-connected space that covers $SO(3)$. Because the fundamental group of $SO(3)$ is $\mathbb{Z}_2$, this space will be a "double cover." What is this double cover? Why, it's $SU(2)$!

It's not hard to see why this is. Every matrix in $SU(2)$ can be expressed as

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

where $|\alpha|^2 + |\beta|^2 = 1$. If we express these complex variables as $\alpha = x + iy$ and $\beta = z + iw$, then this condition is just $x^2 + y^2 + z^2 + w^2 = 1$. In other words, $SU(2)$ is the same as S^3, the 3-sphere!

Everybody knows that S^3 is simply connected. (S^2, which is much easier to picture, is also simply connected, if that helps.) The fact that S^3 is the universal cover of $\mathbb{R}P^3$ is a basic topological fact which has silently and inevitably been guiding our representation theory this whole time. This double cover can be seen explicitly if we take π to be our covering map. We say that for any matrix $g \in SU(2)$, $\pi(g) = \pi(-g)$. Therefore π is a 2-to-1 map!

You can also see the double-cover manifested in the example we did above.

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \xrightarrow{\pi} \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This fact can be used to send a loop in $SU(2)$ to a loop in $SO(3)$. If we start θ at 0 and slide it to 2π, the $SU(2)$ element will start and end at $\mathbf{1}$. The $SO(3)$ element will also start and end at $\mathbf{1}$, but it will also visit $\mathbf{1}$ at the intermediate value of $\theta = \pi$. One loop in $SU(2)$ projects down to a double loop in $SO(3)$. As S^3 is simply connected, this loop can be contracted. The projection of this contraction will be a contraction of the double loop in $SO(3)$.

You might be wondering if it's possible to make a homomorphism from $SU(2)$ to $SO(3)$ that *doesn't* have this property. In the example above it looks so simple! Just replace 2θ with θ! The problem is that, while you can do this for any one $U(1)$ subgroup of $SU(2)$, you can't do it consistently for the whole group. There are two ways to see this. The first is that we *proved*, using representation theory, that all irreducible representation of $SU(2)$ into 3×3 matrices must be equivalent. The second is topological: you can't have an isomorphism between two spaces with different fundamental groups.

How strange this is. $SO(3)$ is usually thought of the group of rotational symmetries of space. That seems abstract enough, if you ask me! But in order to expose its true nature, we had consider $SO(3)$ *itself* as a space. This allowed us, first of all, to study its tangent space $\mathfrak{so}(3)$. However, this also let us study the global structure of $SO(3)$ using algebraic topology, by looking at the fundamental group *of a group*!

The fact that $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic lets us use our classification of irreps of $SU(2)$ to classify the irreps of $SO(3)$. The result is that all of the integer spin representations of $SU(2)$ can also be regarded as representations of $SO(3)$. In such cases, we usually refer to the integer spin "j" as "l".

There's another thing I should mention: In general, the universal cover of $SO(n)$ is a group called $Spin(n)$. Obviously, $Spin(3)$ is isomorphic to $SU(2)$.

Let's get back to the fun stuff. Let's talk about the Belt Trick!

The Belt Trick is one of those things that's often shown but rarely explained. It's a way of physically demonstrating that the fundamental group of $SO(3)$ is $\mathbb{Z}_2$. In order to understand it, we'll need to discuss another way to picture elements of $SO(3)$.

A system of coordinates in $\mathbb{R}^3$ can be uniquely specified by three ordered orthogonal unit vectors, $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$. There is a unique rotation that will bring the standard basis $\{\hat{x}, \hat{y}, \hat{z}\}$ into this configuration. Therefore, we can uniquely associate each rotation with a set of three ordered orthonormal vectors.

A loop in $SO(3)$ parameterized by $t \in [0,1]$ will start and end at the same element of $SO(3)$. We can represent such a path in $SO(3)$ using a belt. At every point along the belt we have one vector pointing along the length of the belt, one pointing across the width of the belt, and one pointing "out" of the belt. Therefore, each point along the belt specifies a set of three ordered orthonormal vectors, and therefore a unique rotation. As long as one end of the belt is parallel to the other end of the belt, this path through $SO(3)$ will be a loop. This is shown in Figure 10.

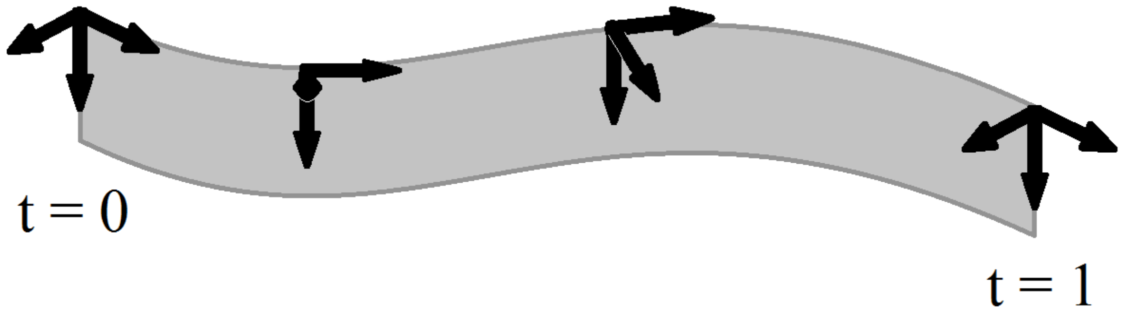

Figure 10: A belt can be used to represent a unique path in $SO(3)$. At any point along the belt, there are three associated vectors: the vector along the belt, the vector pointing across the width of the belt, and the vector pointing out of the belt. If both ends of the belt are parallel, this can be read as a unique loop in $SO(3)$, running from $t = 0$ to $t = 1$.

If the belt is flat like a board, it corresponds to the trivial loop. If the one end of the belt is twisted by 2π, it corresponds to the non-contractible loop. If the belt is twisted by 4π, it corresponds to going around the non-contractible loop twice (which we know can be deformed into the trivial loop). The actual "trick" is to show that the belt with a 2π twist in it can't be deformed into the flat belt while keeping the ends of the belt parallel the whole time. The belt with a 4π twist, however, can. This is done in Figure 11.

Figure 11: The Belt Trick: A belt with a 4π twist in it can be made flat without rotating the end. The book is there to keep the other end of the belt flat.

16 What's The Deal With Spin?

The strangest aspects of quantum mechanics are just physical manifestations of linear algebra, spin included.

Even once you take quantum mechanics and figure out *what* spin is, you still wonder *why* spin is. Here is the reason: spin *does* exist because spin *can* exist.

Let's approach this axiomatically.

Every quantum system is modeled by a Hilbert space $\mathcal{H}$. The simplest Hilbert spaces are just the vector spaces $\mathbb{C}^n$. Let's assume we have such a Hilbert space, because its certainty allowed by quantum mechanics! Let's assume something else: let's assume that our state vectors transform as a unitary irreducible representation of $SO(3)$.

Why should we assume this? Well, in real life we can perform $SO(3)$ rotations by changing reference frames. If we have observables (self adjoint operators on $\mathcal{H}$) that are determined by our reference frame, then the state we would use to describe our system should change after a rotation. However, assuming that no reference frame is preferred in nature, $SO(3)$ must act on our state as a group action. This action must be linear in order to respect the principle of superposition. A linear group action on a vector space is just a representation. Furthermore, we are only interested in *irreducible* representations because all *reducible* just ones break up into irreducible ones anyway. Therefore we can save ourselves time and just consider the irreducible ones. Finally, this representation should be unitary because it should preserve the inner product of $\mathcal{H}$, implying that we can't detect through measurements that our reference frame has rotated.

There is one unitary irreducible representation of $SO(3)$ for each integer l. The dimension of that irrep is $2l+1$. Therefore, we conclude that $\mathcal{H} = \mathbb{C}^n$ can transform irreducibly under $SO(3)$ only when n is odd!

Wait... We must have screwed something up.

Hm.

What could it be...

...!

Aha!

We wanted our group action on $\mathcal{H}$ to be a unitary representation because we wanted the action of our symmetry group to not affect what we could observe. However, the overall *phase* of a quantum state is never observable! So really, we were being too strict. In order to ensure that the action of our symmetry group does not affect anything we could measure, we only need to ensure that our Hilbert space transforms as a unitary representation of $SO(3)$ only *up to a phase*!

Let's try to make a map from $SO(3)$ to its universal cover $SU(2)$. We already know that we have a 2-to-1 homomorphism $\pi : SU(2) \to SO(3)$. We can try to "invert" π, but it won't be perfect because π isn't a bijection. In any case, we can create a map $\rho : SO(3) \to SU(2)$ that sends an element of $SO(3)$ to one of the two $SU(2)$ matrices in the pre-image. In other words, $\pi \circ \rho = \mathbf{1}$, but $\rho \circ \pi = \pm \mathbf{1}$.

This map ρ is what some would call a "projective representation." A projective representation is a lot like a representation, except that its not quite a homomorphism. The identity is mapped to a phase

$$\rho(\mathbf{1}) = e^{i\theta}\mathbf{1}$$

and the homomorphism is off by a phase that depends on the two group elements in the map:

$$\rho(g)\rho(h) = e^{i\theta(g,h)}\rho(gh).$$

Note that the phase, called the "cocycle," automatically satisfies the "cocycle relation"

$$e^{i\theta(g,h)}e^{i\theta(gh,k)} = e^{i\theta(g,hk)}e^{i\theta(h,k)}.$$

Of course, for the ρ we defined above, our phases are all ± 1, but its nice to know the general definition. Certainly, we can always change the definition of our projective representation so the phases are different:

$$\rho'(g) = e^{i\varphi(g)}\rho(g).$$

We would then have

$$\rho'(1) = e^{i\theta'}\mathbf{1}$$
$$\rho'(g)\rho'(h) = e^{i\theta'(g,h)}\rho'(gh)$$

where

$$\theta' = \theta + \varphi(1)$$
$$\theta'(g,h) = \theta(g,h) + \varphi(g) + \varphi(h) - \varphi(gh)$$

We can pretty clearly always choose a φ so $\theta' = 0$, but we can not always change it so $\theta(g,h) = 0$. If this is the case, we say we have an "essential cocycle." For example, the cocycle in our map from $SO(3)$ into $SU(2)$ is essential. This may seem suspiciously cohomological to you. It turns out that if a Lie group is simply connected (and the "central charge" of the Lie algebra can be removed) then no cocycle is essential. Weinberg proves this in chapter 2, appendix B of *The Quantum Theory of Fields, Vol 1*, which I haven't read.

Let's return to earth. For the n dimensional irreducible representation π_n of $SU(2)$,

$$\pi_n : SU(2) \to GL(\mathbb{C}^n).$$

we have a projective representation of $SO(3)$ into $GL(\mathbb{C}^n)$

$$\rho_n : SO(3) \to GL(\mathbb{C}^n)$$

defined by

$$\rho_n(g) \equiv \pi_n \circ \rho.$$

We can see that ρ_n is a projective representation through the homomorphism property of π_n.

$$\begin{aligned}\rho_n(gh) &= \pi_n\big(e^{-i\theta(g,h)}\rho(g)\rho(h)\big)\\ &= \pi_n(e^{-i\theta(g,h)}\mathbf{1})\rho(g)\rho(h)\end{aligned}$$

Because $e^{-i\theta(g,h)}\mathbf{1}$ commutes with all of $SU(2)$, $\pi_n(e^{-i\theta(g,h)}\mathbf{1})$ commutes with all $\pi_n(g)$. As π_n is an irreducible representation, by Schur's Lemma we know that

$\pi_n(e^{-i\theta(g,h)}\mathbf{1})$ must be a constant. As we know that the determinant must be 1, we know that $\pi_n(e^{-i\theta(g,h)}\mathbf{1})$ must be $\mathbf{1}$ if n is odd and $\pm\mathbf{1}$ if n is even. Therefore, ρ_n is a true representation of $SO(3)$ if n is odd, but may only be (and actually is) a projective representation if n is even.

We have classified all of the unitary irreducible representations of $SU(2)$: there is one for each n. Therefore, there is a *projective* representation of $SO(3)$ for each n. Apparently, in order to classify the projective representations of a group, we need to look at its universal cover!

Furthermore, we know that these are the *only* irreducible projective representations of $SO(3)$. Say there were an irreducible projective representation of $SO(3)$ with an essential cocycle that wasn't born from an $SU(2)$ representation. We could then make a projective representation of $SU(2)$ with an essential cocycle, which doesn't exist.

By the way, all this "projective representation" stuff is what people are talking about when they say that you have to "rotate" an electron twice in order to get back to where you started. They just mean that $SO(3)$ can only be represented on $\mathbb{C}^2$ projectively. If you try to lift the the non-contractible loop in $SO(3)$ to $SU(2)$ in a continuous way, the lifted loop will start at $\mathbf{1}$ but end at $-\mathbf{1}$.

The projective representations of $SO(3)$ are the half-integer spin representations of $SU(2)$. These are the spins that we missed before we considered projective representations.

The reason why we *do* observe half-integer spins in nature is because there's no reason why we shouldn't be able to! A projective representation of $SO(3)$ is just as valid as an actual representation.

This, I hope you agree, is all very bizarre. A strange fact about the space of rotations admits a loop-hole in quantum mechanics that allows for some very strange objects. The fact that Nature actually does realize these trippy doodads just goes to show how much more creative she is than we. In fact, she didn't only realize them: she made a whole universe out of them! Your entire body is made out of spin $\frac{1}{2}$ particles!

But let's not get carried away. As we have seen, even Nature is constrained by both topology and representation theory. Sure, we found some curious objects, but we also found *all of them*. There is no particle with spin $\frac{1}{3}$ or spin $\frac{5}{6}$, for example. Just half-integers. It's very peculiar: Nature can be creative, but only within some very rigid (and fundamentally tautological) bounds. It's enough to give you that ancient platonic feeling, albeit with a modern spooky twist.

17 The Adjoint Representation And The Power Of Magical Thinking

An quantum mechanical "observable" $\hat{O}$ is just a self-adjoint operator on a Hilbert space $\mathcal{H}$.

$$\hat{O} : \mathcal{H} \to \mathcal{H} \qquad\qquad \hat{O}^\dagger = \hat{O}$$

Conjugation by a unitary operator U will leave our operator self adjoint. That is,

$$(U\hat{O}U^\dagger)^\dagger = U\hat{O}U^\dagger.$$

Similarly, conjugation of a skew-adjoint operator by U will leave the operator skew-adjoint. (As every self-adjoint operator is just i times a skew-adjoint operator, we can see that these two facts are equivalent.)

This phenomenon is a special case of more general fact, one that we have already proved. For a Lie group G with a Lie algebra $\mathfrak{g}$,

$$gXg^{-1} \in \mathfrak{g} \text{ for all } g \in G, X \in \mathfrak{g}.$$

What's more, for any homomorphism $\pi : G_1 \to G_2$ with a Lie algebra representation $\pi' : \mathfrak{g}_1 \to \mathfrak{g}_2$, we proved

$$\pi(g)\pi'(X)\pi(g)^{-1} \in \mathfrak{g}_2 \text{ for all } g \in G_1, X \in \mathfrak{g}_1.$$

$U(n)$, the group of $n \times n$ unitary matrices, has the Lie algebra $\mathfrak{u}(n)$, the Lie algebra of $n \times n$ skew-adjoint matrices. Therefore, for a unitary representation of G,

$$\pi : G \to U(n)$$

we have

$$\pi(g)\pi'(X)\pi(g)^{-1} \in \mathfrak{u}(n) \text{ for all } g \in G, X \in \mathfrak{g}.$$

$\pi'(X)$ is skew-adoint, so $i\pi'(X)$ is a self adjoint operator on $\mathbb{C}^n$. It is therefore an observable.

Never forget that Lie algebras are just real vector spaces (that are closed under a Lie bracket). We can pretty clearly see that we have a group action of G on $\mathfrak{g}$, defined by conjugation. As $\mathfrak{g}$ is a vector space, we have a group homomorphism into $GL(\mathfrak{g})$. We denote this homomorphism by "*Ad*."

$$Ad : G \to GL(\mathfrak{g})$$
$$Ad(g)X \equiv gXg^{-1}$$

(Note that if we defined the action as $g^{-1}Xg$ instead of gXg^{-1} we would have a right action instead of a left one.) We can confirm that Ad truly is a representation by checking that its action is linear.

$$\begin{aligned} Ad(g)(X+Y) &= g(X+Y)g^{-1} \\ &= gXg^{-1} + gYg^{-1} \\ &= A(g)X + Ad(g)Y \quad \checkmark \end{aligned}$$

We call Ad the "Adjoint representation" of G.

The existence of the Adjoint representation, combined our with our observation that observables come from the Lie algebras of unitary representations, gives us a new way to study what happens when we have a group representation on a Hilbert space.

Let's take $SO(3)$, for example. One of the fundamental principles of quantum mechanics is that information can only be extracted from a state via observation. If we have a Hilbert space $\mathcal{H}$ that "lives" in three dimensional space, it will transform as a (projective) unitary representation of $SO(3)$. A reference frame in three dimensional space is given by three orthonormal vectors, $\{\hat{x}, \hat{y}, \hat{z}\}$. The only way that we could "tell" if our state has been rotated is if we have three observables

$\{\hat{O}_x, \hat{O}_y, \hat{O}_z\}$, one for each orthonormal vector, and use them to measure our state. Somehow, our $SO(3)$ representation must be encoded in these operators. For example, a rotation around the x-axis should not change the action of the $\hat{O}_x$ observable. Magical thinking suggests that these operators should somehow be "derivable" from our representation.

Now, because there is fundamentally no difference between our unit vectors except for their relationships to each other, it would be nice is these observables depended solely on the unit vector itself:

$$\hat{O}_x = \hat{O}(\hat{x}) \qquad \hat{O}_y = \hat{O}(\hat{y}) \qquad \hat{O}_z = \hat{O}(\hat{z})$$

A rotation of space is given by a matrix $R \in SO(3)$ and acts on our unit vectors in the standard way, and acts on $\mathcal{H}$ as a unitary projective representation. Say that the unitary matrix that corresponds to R is "$U(R)$." The key observation is that, if $SO(3)$ is truly a symmetry of space, then if we rotate one of these observables *as well as* our state, we should never be able to measure the effect of this rotation. In other words, for all unit vectors $\hat{n}$, we want

$$\langle\phi| U(R)^\dagger \hat{O}(R\hat{n}) U(R) |\psi\rangle = \langle\phi|\psi\rangle .$$

This only holds if

$$\hat{O}(R\hat{n}) = U(R)\hat{O}(\hat{n})U(R)^{-1}.$$

(Note that even if U is a projective representation and not a true representation, the phase will be canceled out by $U(R)^{-1}$ in the expression above.) Therefore, we can see see that our observables $\hat{O}(\hat{n})$ will succeed at measuring our states in a way that respects the rotational symmetry of space only if they transform in a particular way, dictated by our (projective) representation.

That's all fine and good. But what if you don't know what $\hat{O}(\hat{n})$ is? What if you know how $\mathcal{H}$ transforms under rotations, but you *don't know* what observables are up to the task? How can you find $\hat{O}(\hat{n})$ that satisfies the condition above?

Well, seeing as these observables somehow "capture" all of the available information of this group action, it seems as though there should be some connection between them and the representation. Lo and behold, you can just use (i times) the operators in the Lie algebra of you unitary representation, which readily transform under the Adjoint representation! This seems promising, so let's see if it works.

As we saw, projective representations of $SO(3)$ are just actual representations of its universal cover $SU(2)$. Therefore, it suffices to consider all of the irreps of $SU(2)$.

Say $SU(2)$ acts on our Hilbert space $\mathbb{C}^n$ as a unitary irrep.

$$\pi : SU(2) \to U(n)$$

The Lie algebra of $\mathfrak{su}(2)$ is spanned by the skew-adjoint matrices X_j. (They correspond to the generators L_j of $\mathfrak{so}(3)$.) We can then define the self-adjoint operator $\hat{O}(\hat{n})$ as follows

$$\hat{O}(\hat{n}) \equiv i\pi'(\hat{n} \cdot \vec{X})$$

For the spin $\frac{1}{2}$ representation of $SU(2)$, these matrices will be

$$\hat{O}(\hat{n}) = \tfrac{1}{2}\hat{n}\cdot\vec{\sigma}.$$

This operator will satisfy the transformation property we want. Let's check this explicitly. Recall that, for either of the two $g \in SU(2)$ that correspond to $R \in SO(3)$, we have

$$g(\vec{x}\cdot\vec{\sigma})g^{-1} = (R\vec{x})\cdot\vec{\sigma}.$$

Our transformation property follows from the fact that $\vec{X} \propto \vec{\sigma}$.

$$\begin{aligned}
\pi(g)\hat{O}(\hat{n})\pi(g)^{-1} &= \pi(g)i\pi'(\hat{n}\cdot\vec{X})\pi(g)^{-1} \\
&= i\pi'\big(g(\hat{n}\cdot\vec{X})g^{-1}\big) \\
&= i\pi'\big((R\hat{n})\cdot\vec{X}\big) \\
&= \hat{O}(R\hat{n}) \quad \checkmark
\end{aligned}$$

We have therefore constructed a set of operators that will rotate properly for every projective representation of $SO(3)$. I hope you recognize that these are the spin operators, often denoted $\hat{S}_x, \hat{S}_y, \hat{S}_z$.

We've done a lot, so let's review the philosophy behind our math: if a Hilbert space transforms under a projective unitary representation of a group, that group can rightfully be regarded as a group of "symmetries" of our Hilbert space because it preserves the inner product. Because quantum states only convey information insofar as they can be measured, we want to find a set of operators that transform in an interesting way along with our group. Such operators are automatically given by $\pm i\pi'(X)$ for $X \in \mathfrak{g}$, where X transforms under G according to the "Adjoint representation" given by conjugation. Furthermore, we note that the observable corresponding to the Lie algebra element X will be unaffected by the group transformation that X generates:

$$\begin{aligned}
\pi(e^{tX})i\pi'(X)\pi(e^{-tX}) &= i\pi'(e^{tX}Xe^{-tX}) \\
&= i\pi'(X)
\end{aligned}$$

Wow! That sure is a lot of interesting structure! Let's look at more examples of observables coming from Lie algebra representations!

States of a spin-less particle in one dimension live in the Hilbert space $\mathcal{H} = L^2$, where L^2 is the space of square integrable functions from $\mathbb{R} \to \mathbb{C}$. This Hilbert space has an inner product given by

$$\langle f, g\rangle \equiv \int \mathrm{d}x\, f^*(x)g(x).$$

Just by eyeballing this inner product, we see two unitary group representations π_1 and π_2 of the group $\mathbb{R}$. The first one acts as

$$\big(\pi_1(a)f\big)(x) = f(x+a)$$

and the second acts as

$$\big(\pi_2(b)f\big)(x) = e^{ibx}f(x)$$

where $a, b \in \mathbb{R}$. Its easy to check that this is a representation of the additive group $\mathbb{R}$. The Lie algebra elements of $\mathbb{R}$ under this representation are obtained from the infinitesimal action of π_1 and π_2.

$$\begin{aligned}
\Big(\lim_{\epsilon\to 0}\frac{1}{\epsilon}\big(\pi_1(\epsilon)-\pi_1(0)\big)f\Big)(x) &= \lim_{\epsilon\to 0}\frac{f(x+\epsilon)-f(x)}{\epsilon} \\
&= \frac{\mathrm{d}}{\mathrm{d}x}f(x)
\end{aligned}$$

$$\begin{aligned}
\Big(\lim_{\epsilon\to 0}\frac{1}{\epsilon}\big(\pi_2(\epsilon)-\pi_2(0)\big)f\Big)(x) &= \lim_{\epsilon\to 0}\frac{e^{i\epsilon x}f(x)-f(x)}{\epsilon} \\
&= ixf(x)
\end{aligned}$$

As expected, these operators are skew-adjoint. Multiplication by $-i$ will yield our two favorite observables, obtained directly from our unitary representations:

$$\begin{aligned}
\pi_1 &\longrightarrow -i\frac{\mathrm{d}}{\mathrm{d}x} \\
\pi_2 &\longrightarrow \quad x
\end{aligned}$$

Let's see how these operators act under conjugation by their respective unitary representations. Because $\mathbb{R}$ is abelian, and our operators are infinitesimal elements of our representations, they will commute with elements of the representations. Using this fact, we can see that, under conjugation, these operators remain invariant. (This means the Adjoint representation is trivial, as is always the case for an abelian Lie group.)

$$\begin{aligned}
\pi_1(a)\Big(-i\frac{\mathrm{d}}{\mathrm{d}x}\Big)\pi_1(a)^{-1} &= \pi_1(a)\pi_1(a)^{-1}\Big(-i\frac{\mathrm{d}}{\mathrm{d}x}\Big) \\
&= -i\frac{\mathrm{d}}{\mathrm{d}x}
\end{aligned}$$

$$\begin{aligned}
\pi_2(b)\big(x\big)\pi_2(b)^{-1} &= \pi_2(b)\pi_2(b)^{-1}\big(x\big) \\
&= x
\end{aligned}$$

This makes sense once you think about it. π_1 is the group of translations in position space and π_2 is the group of translations in momentum space. A momentum measurement will be unaffected by a translation in position space and a position measurement will be unaffected by a translation in momentum space. Duh.

This is should be clear because each representation is generated by these Lie algebra elements!

$$\begin{aligned}
\big(e^{a\frac{\mathrm{d}}{\mathrm{d}x}}f\big)(x) &= \sum_{n=0}^{\infty}\frac{a^n}{n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}f(x) \\
&= f(x+a) \\
&= \big(\pi_1(a)f\big)(x)
\end{aligned}$$

$$\left(e^{ibx} f\right)(x) = \left(\pi_2(b) f\right)(x)$$

This is therefore just a specific example of the observable $i\pi(X)$ not changing from conjugation by $\pi(e^{tX})$.

Let's do one last example!

States of a spin-less particle in three dimensions are square integrable functions from $\mathbb{R}^3 \to \mathbb{C}$. The inner product of $\mathcal{H}$ is given by

$$\langle f, g\rangle \equiv \int \mathrm{d}^3 x f^*(\vec{x}) g(\vec{x}).$$

We still have our position-space and momentum-space transformations that we investigates last time. They will give the position and momentum operators in each of the three directions. There is a more exciting unitary representation that acts on this space: $SO(3)$ (again). The group action is the obvious one.

$$\left(\pi(R) f\right)(\vec{x}) = f(R^{-1}\vec{x})$$

Its unitarity can be checked with a change of variables.

Recall that the basis of $\mathfrak{so}(3)$ is given by

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad L_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad L_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Unlike $\mathbb{R}$, $SO(3)$ is non-abelian, so the Adjoint representation will be a bit juicier. Therefore, once we find the Lie algebra elements of this unitary representation, we expect to get a set of observables that transform non-trivially under rotations.

Let's find those observables. Once again, they'll be given by elements of our Lie algebra homomorphism. Lets now figure out what the operators $\pi'(L_j)$ are. We'll start with $\pi'(L_1)$. Using the relation

$$\pi(e^{\theta L_1}) = e^{\theta \pi'(L_1)}$$

we have

$$\begin{aligned}
\left(\pi'(L_1) f\right)(\vec{x}) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(e^{\theta\pi'(L_1)} f\right)(\vec{x})\bigg|_{\theta=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\pi(e^{\theta L_1}) f\right)(\vec{x})\bigg|_{\theta=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\, f(e^{-\theta L_1}\vec{x})\bigg|_{\theta=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\, f(x, \cos\theta y + \sin\theta z, -\sin\theta y + \cos\theta z)\bigg|_{\theta=0} \\
&= \left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right) f(\vec{x})
\end{aligned}$$

which follows from the chain rule. Doing this for L_2 and L_3 as well, we find

$$\pi'(L_1) = \left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right) \qquad \pi'(L_2) = \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \qquad \pi'(L_3) = \left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right).$$

Multiplication by $-i$ gives us the self-adjoint angular momentum operators! Furthermore, we can see that the angular momentum operator in the $\hat{n}$ direction, $-i\hat{n}\cdot\pi'(\vec{L})$, generates rotations around the $\hat{n}$ axis! Apparently, if you rotate a state around an axis, the angular momentum component in that direction will not change.

Intuitively, we expect that conjugation by a rotation matrix will rotate the component of angular momentum we are measuring. In other words, we expect for the following equation to hold:

$$\pi(R)\pi'(\hat{n}\cdot\vec{L})\pi(R)^{-1} = \pi'\big((R\hat{n})\cdot\vec{L}\big).$$

Is this the case? Indeed it is. It is a consequence of the fact that

$$R(\hat{n}\cdot\vec{L})R^{-1} = (R\hat{n})\cdot\vec{L}$$

which I encourage you to check for yourself. This also implies that the Adjoint representation of $SO(3)$ is just the defining representation of $SO(3)$!

This should all shed some light on why spin and angular momentum are so similar. In both cases, there's a Hilbert space with a (possibly projective) unitary representation of $SO(3)$ acting on it. The spin and angular momentum operators are both born from the Lie algebra of $SO(3)$ (or its double cover $SU(2)$) and thus behave in the same way under rotation!

Let's take a step back reflect on what we've accomplished. It seems as though we've done so much that its difficult to figure out *what* exactly we did! We have the embarrassment of riches that so often characterizes representation theory!

"Symmetry" can mean a lot of things. Often, when physicists say they have a "symmetry" of a system, they mean they have a group of transformations that doesn't affect the Hamiltonian (or Lagrangian). However, in this section, we did not *once* mention the Hamiltonian, or assumed that something commuted with the Hamiltonian, or anything like that at all! Time evolution was never once brought into the picture.

Instead, when we said we "had a symmetry," we meant that there was a unitary representation that acted on our space. This is really far more fundamental than an operator that commutes with the Hamiltonian. The mere *presence* of a unitary representation acting on our Hilbert space is already interesting enough as is. The Lie algebra elements of these unitary representations both *generate* our representation and *transform* under it. The Lie algebra, of course, transforms under the group via conjugation, i.e. under the Adjoint representation. Therefore, the observables that come from the Lie algebra of our unitary representation will transform in an interesting way under the group. It does, of course, take initiative to even ask the fundamental question motivating this all:

> What are the interesting things I can measure from my Hilbert space?
>
> —You (hopefully)

In non-relativistic quantum mechanics there are only so many observables to talk about: spin, position, momentum, angular momentum. Clearly, when we analyze the simplest Hilbert spaces we have, we're just going to rediscover the operators we already knew about. However, even then, we've still shed so much light on their true nature by unearthing them in this way. If it so happens that the operators we find commute with the Hamiltonian, well, then we have all the more reason to celebrate!

But even if they don't, we've still discovered something of incredible worth. After all, physics is just the friends you make along the way.

18 Operators That Generate Representations of $U(1)$ Have Quantized Eigenvalues

"Quantum." We say that word so much that we forget what it means. Never forget: a "quantum" is a discrete amount. The process of "quantization," going from coordinates in a phase space to operators on a Hilbert space, is a confusing name! Doesn't it seem as though the classical system is actually *more* "quantum" than the quantum system? I mean, a set of point-like particles all traveling along well-defined paths seems more discrete than a smudgy wave-function.

No... what puts the "quantum" in "quantum mechanics" is that *certain quantities* can only be measured to have discrete values. Spin and angular momentum are prime examples of this. Having said that, there are of course *other* quantities that *aren't* discrete. For example, a particle moving in one dimension can be measured to be at any position or have any momentum. These observables don't have quantized eigenvalues, while spin and angular momentum do. What gives?

The answer is simple: only the operators that generate $U(1)$ representations have quantized eigenvalues.

Say that you have an observable $\hat{O} : \mathcal{H} \to \mathcal{H}$. Multiplication by it (for some real number t) will yield a skew-adjoint matrix which can then be exponentiated to yield a unitary matrix.

$$\hat{O} \longrightarrow U(t) = e^{it\hat{O}}$$

This is clearly a unitary representation of $\mathbb{R}$:

$$\begin{aligned} U(t_1)U(t_2) &= e^{it_1\hat{O}}e^{it_2\hat{O}} \\ &= e^{i(t_1+t_2)\hat{O}} \\ &= U(t_1 + t_2). \end{aligned}$$

For operators like position and momentum, the story ends here. However, *sometimes*, we'll get another extra little gift. Sometimes, there will be a T where

$$U(2\pi T) = 1.$$

Now we're in business.

This allows us to create a $U(1)$ representation on our state space, defined in the obvious way:

$$U(e^{i\theta}) = e^{i\theta T\hat{O}}.$$

But we've seen this story before! The Hilbert space $\mathcal{H}$ will just break up into irreducible representations of this $U(1)$ action, and $\hat{O}$ will be of the form

$$\hat{O} = \frac{1}{T}\begin{pmatrix} k_1 & 0 & 0 & \cdots \\ 0 & k_2 & 0 & \cdots \\ 0 & 0 & k_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

for some collection of integers k_i! Therefore, the values of $\hat{O}$ will be quantized in units of $\frac{1}{T}$!

Let's look at angular momentum. Using the notation from the last part, the angular momentum in the $\hat{n}$ direction is given by

$$-i\hbar\pi'(\hat{n} \cdot \vec{L})$$

(properly including the $\hbar$).

Multiplying by it and exponentiating, the unitary representation this operator generates is

$$U(t) = e^{t\hbar\pi'(\hat{n}\cdot\vec{L})}$$

But we know something: the matrix

$$e^{\theta\hat{n}\cdot\vec{L}}$$

is the rotation matrix in $SO(3)$ around the axis $\hat{n}$. Therefore, when $\theta = 2\pi$ we have

$$e^{(2\pi)\hat{n}\cdot\vec{L}} = 1$$

and that's all we need, as

$$e^{t\hbar\pi'(\hat{n}\cdot\vec{L})} = \pi\left(e^{t\hbar\hat{n}\cdot\vec{L}}\right).$$

We can see that our "T," the number for which $U(2\pi T) = 1$, is

$$T = \frac{1}{\hbar}.$$

Therefore, angular momentum is quantized in units of $\hbar$!

Maybe its not surprising that spin and angular momentum are quantized, while position and momentum are not. After all, thinking about the previous section, the spin and angular momentum operators are born from representations of $SU(2)$ and $SO(3)$, respectively. Indeed, both groups are groups of "rotations," and thus have a lot of $U(1)$'s in them! (By the way, this also applies to electric charge. In that case, the $U(1)$ representation acts as a rotation in the ϕ-ϕ^* plane of a complex-valued field.)

There is another famous observable with "quantized" eigenvalues, namely the Hamiltonian of the harmonic oscillator. Usually, this is seen by using the commutation relations of the creation and annihilation operators. That's a perfectly good way to do things, but it obscures "why" the energy levels are quantized. I will now present a fun little proof I came up with that explicitly draws the connection between $U(1)$ and the harmonic oscillator.

I haven't yet discussed canonical quantization yet, but the idea is simple. The coordinates q and p evolve in in time as

$$\frac{\mathrm{d}}{\mathrm{d}t}q = \{q, H\} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}p = \{p, H\}$$

and the Heisenberg operators $\hat{q}$ and $\hat{p}$ evolve in time as

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{q} = -\frac{i}{\hbar}[\hat{q}, \hat{H}] \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}\hat{p} = -\frac{i}{\hbar}[\hat{p}, \hat{H}].$$

$\hat{H}$ comes from H by just replacing all of the q's and p's with $\hat{q}$ and $\hat{p}$. (This cannot always be done consistently, particularly if there are operator ordering ambiguities, but that is unimportant for the case at hand.)

The classical Hamiltonian of the 1-dimensional harmonic oscillator is

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

and thus the Hamiltonian of corresponding quantum system is

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{m\omega^2}{2}\hat{q}^2.$$

In classical mechanics, these differential equations are easy to solve. Points in phase space will move around the origin in a circle with frequency ω.

$$\frac{\mathrm{d}}{\mathrm{d}t}q = \frac{1}{m}p \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}p = -m\omega^2 q$$

$$\Longrightarrow q(t) = q(0)\cos(\omega t) + \frac{1}{m\omega}p(0)\sin(\omega t) \qquad\qquad p(t) = p(0)\cos(\omega t) - m\omega q(0)\sin(\omega t)$$

At each interval of $2\pi/\omega$, every point in phase space will return to its starting point. (You knew this in high school: the frequency of a pendulum does not depend on the amplitude of its oscillation.)

$$q(\tfrac{2\pi}{\omega}n) = q(0) \qquad\qquad p(\tfrac{2\pi}{\omega}n) = p(0) \qquad \text{for all } n \in \mathbb{Z}$$

Moving up to the operator equation, using the canonical commutation relation $[\hat{q}, \hat{p}] = i\hbar$, we find that the differential equations given by time evolution are completely identical.

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{q} = \frac{1}{m}\hat{p} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}\hat{p} = -m\omega^2 \hat{q}$$

$$\Longrightarrow \hat{q}(t) = \hat{q}(0)\cos(\omega t) + \frac{1}{m\omega}\hat{p}(0)\sin(\omega t) \qquad\qquad \hat{p}(t) = \hat{p}(0)\cos(\omega t) - m\omega \hat{q}(0)\sin(\omega t)$$

Just as before,

$$\hat{q}(\tfrac{2\pi}{\omega}n) = \hat{q}(0) \qquad\qquad \hat{p}(\tfrac{2\pi}{\omega}n) = \hat{p}(0) \qquad \text{for all } n \in \mathbb{Z}$$

Recall that operators evolve in time via conjugation by the time evolution operator.

$$U(t) = e^{-it\hat{H}/\hbar}$$

$$\hat{q}(t) = U(-t)\hat{q}(0)U(t) \qquad\qquad \hat{p}(t) = U(-t)\hat{p}(0)U(t)$$

The states $|q\rangle$ are eigenstates of $\hat{q}(0)$.

$$\hat{q}(0)\,|q\rangle = q\,|q\rangle$$

That means that the states $|q\rangle$ are also eigenstates of $\hat{q}(t)$ when $t = \frac{2\pi}{\omega}n$.

$$\hat{q}(\tfrac{2\pi}{\omega}n)\,|q\rangle = \hat{q}(0)\,|q\rangle = q\,|q\rangle$$

Therefore, $U(t)\,|q\rangle$ is an eigenstate of $\hat{q}(0)$:

$$U(-t)\hat{q}(0)U(t)\,|q\rangle = q\,|q\rangle \implies \hat{q}(0)U(t)\,|q\rangle = qU(t)\,|q\rangle \qquad \text{for } t = \tfrac{2\pi}{\omega}n$$

As the eigenspaces of the position operator $\hat{q}(0)$ are one dimensional, this means that $U(\frac{2\pi}{\omega}n)\,|q\rangle$ is proportional to $|q\rangle$. Furthermore, as $U(\frac{2\pi}{\omega}n)$ is unitary, this proportionality factor can only be (position dependent) phase.

$$U(\tfrac{2\pi}{\omega}n)\,|q\rangle = e^{in\theta(q)}\,|q\rangle$$

We could make the exact same argument for the momentum operator.

$$U(\tfrac{2\pi}{\omega}n)\,|p\rangle = e^{in\varphi(p)}\,|p\rangle$$

As the position eigenstates and the momentum eigenstates both comprise a complete basis of the Hilbert space, this implies that

$$e^{i\theta(\hat{q})} = e^{i\varphi(\hat{p})}.$$

We can act with both operators on the wave function $\psi(q) = 1$. As $\hat{p}\,|\psi\rangle = 0$, this implies that

$$e^{i\theta(q)} = e^{i\varphi(0)}$$

which implies that $\theta(q)$ is a constant. Likewise, $\varphi(p)$ must also be a constant. We will call this constant θ.

$$U(\tfrac{2\pi}{\omega}) = e^{i\theta}$$

$e^{i\theta}$ can be set to 1 by shifting the Hamiltonian by a constant term. After the Hamiltonian is shifted and we have

$$U(\tfrac{2\pi}{\omega}) = 1.$$

This is a $U(1)$ representation, and from our previous discussion, the eigenvalues of $\frac{1}{\hbar}\hat{H}$ must come quantized in units of ω!

As a quantum field is just a bunch of harmonic oscillators, we can now see *why* particles come in discrete packets. It all goes back to the irreducible representations of $U(1)$! Apparently, the only thing truly "quantum" about quantum mechanics is $U(1)$, the simplest Lie group.

19 Functions on Phase Space Comprise a Lie Algebra

Let's leave the world of quantum mechanics for a bit and investigate classical mechanics in 1-dimension. "States" are now specified by two real numbers, position and momentum. That is, states are just points in $\mathbb{R}^2$, which we call "phase space" for God knows what reason.

$$(q, p)$$

We can certainly make functions of these variables, which we will often refer to as "functions on phase space."

$$f : (q, p) \to \mathbb{R}$$

Let's now consider a path in phase space parameterized by the variable "t." What are the conditions such that, along this path, the value of $f\big(q(t), p(t)\big)$ doesn't change? (We would say that f is "conserved" along such a trajectory.)

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(q(t), p(t)\big) = 0$$

The chain rule tells us that

$$\frac{\mathrm{d}q}{\mathrm{d}t}\frac{\partial f}{\partial q} + \frac{\mathrm{d}p}{\mathrm{d}t}\frac{\partial f}{\partial p} = 0.$$

Each function $\lambda : \mathbb{R} \to \mathbb{R}$ gives us a solution to this differential equation (assuming we have the initial conditions $q(0)$ and $p(0)$):

$$\big(\frac{\mathrm{d}}{\mathrm{d}t} q(t), \frac{\mathrm{d}}{\mathrm{d}t} p(t)\big) = \Big(\lambda(t)\frac{\partial f}{\partial p}, -\lambda(t)\frac{\partial f}{\partial q}\Big)$$

But $\lambda(t)$ is an unimportant factor which can be removed through a re-parameterization of t. Therefore, the only important solution to this differential equation is

$$\big(\frac{\mathrm{d}}{\mathrm{d}t} q(t), \frac{\mathrm{d}}{\mathrm{d}t} p(t)\big) = \Big(\frac{\partial f}{\partial p}, -\frac{\partial f}{\partial q}\Big)$$

which can be re-written as

$$\frac{\mathrm{d}}{\mathrm{d}t} q = \{q, f\} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t} p = \{p, f\}$$

using the "Poisson bracket" defined below.

$$\{f, g\} \equiv \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial g}{\partial q}\frac{\partial f}{\partial p}$$

Note that the Poisson bracket takes in *functions* on phase space. Something confusing is that "q" and "p" sometimes represent the coordinates of a point in our phase space, and sometimes represent the coordinate functions. Always be sure that you know which one is being used at any given moment.

We have motivated the Poisson bracket as a tool for making trajectories on phase space that conserve a given quantity. However, it has many interesting properties of its own.

Anti-commutativity:

$$\{f, g\} = -\{g, f\}$$

Bilinearity:

$$\{\alpha f_1 + \beta f_2, g\} = \alpha\{f_1, g\} + \beta\{f_2, g\} \text{ for all } \alpha, \beta \in \mathbb{R}$$

The Jacobi Identity:

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$$

In these notes, I have defined Lie algebras as the tangent space of a Lie group. There is, however, a more abstract definition of Lie algebras. A Lie Algebra $\mathfrak{g}$ is often defined to be a real vector space that has a Lie Bracket

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

is that is anti-commutative, bilinear, and satisfies the Jacobi identity. Of course, as we have shown, the tangent space of matrices at the identity of a Lie group, using the commutator as the Lie bracket, satisfies the above conditions. Therefore, our previous definition does not contradict this one. It's just that this definition is more general.

We see that functions on phase space comprise a Lie algebra, with the Lie bracket being the Poisson bracket.

After staring at the Poisson bracket for a moment, we see something interesting. Assuming that f and g are polynomials of q and p, we notice that if the degree of f and g is 2 or less, then the degree of $\{f, g\}$ will also be of degree 2 or less. Therefore, the space of polynomials of degree 2 or less is closed under the Lie bracket. As they form a vector space, they also form a sub Lie algebra. The basis of the sub Lie algebra is given by

$$1 \qquad q \qquad p \qquad \frac{q^2}{2} \qquad \frac{p^2}{2} \qquad qp.$$

The brackets are (excluding all the brackets with 1 that are just 0)

$$\begin{array}{ccccc}
\{q,q\} = 0 & \{p,q\} = -1 & \{\frac{1}{2}q^2, q\} = 0 & \{\frac{1}{2}p^2, q\} = -p & \{qp, q\} = -q \\
\{q,p\} = 1 & \{p,p\} = 0 & \{\frac{1}{2}q^2, p\} = q & \{\frac{1}{2}p^2, p\} = 0 & \{qp, p\} = p \\
\{q, \frac{1}{2}q^2\} = 0 & \{p, \frac{1}{2}q^2\} = -q & \{\frac{1}{2}q^2, \frac{1}{2}q^2\} = 0 & \{\frac{1}{2}p^2, \frac{1}{2}q^2\} = -qp & \{qp, \frac{1}{2}q^2\} = -q^2 \\
\{q, \frac{1}{2}p^2\} = p & \{p, \frac{1}{2}p^2\} = 0 & \{\frac{1}{2}q^2, \frac{1}{2}p^2\} = qp & \{\frac{1}{2}p^2, \frac{1}{2}p^2\} = 0 & \{qp, \frac{1}{2}p^2\} = p^2 \\
\{q, qp\} = q & \{p, qp\} = -p & \{\frac{1}{2}q^2, qp\} = q^2 & \{\frac{1}{2}p^2, qp\} = -p^2 & \{qp, qp\} = 0
\end{array}$$

Any subspace that allows for the polynomials of degree 3 or more must include all polynomials. This is because the Poisson brackets of two polynomial of degree 3 will generically be of degree 4. The Poisson brackets of degree 4 will generically be degree 6, and so on. This is an important fact, so hold on to it.

There is another sub Lie algebra of interest here. It's the one spanned by the three simplest coordinate functions

$$1 \qquad q \qquad p$$

and has only one non-zero Lie bracket:

$$\{q, p\} = 1$$

This simple little Lie algebra is of great importance. It's given an intimidating name: the **Heisenberg Lie algebra**. It turns out that the Heisenberg Lie Algebra can actually be exponentiated into a "Heisenberg group." However, the Heisenberg group is not nearly as important as the Lie algebra.

Let me emphasize how peculiar this all is. We've so far motivated the study of Lie algebras as a way to understands Lie groups that act unitarily on quantum mechanical state spaces. The Lie algebras we've defined here, however, were motivated completely differently. They were defined, not with reference to any Lie group, but only with reference to the fundamental structure of classical phase space, and functions on it.

This goes to show that its tough to come up with a catch-all reason for why representation theory is so useful.

20 The Moment Map: Lie Algebra $\to$ Conserved Quantities

In our section on the Adjoint Representation and Magical Thinking, we found that observables of interest on a Hilbert space are given by Lie algebra elements of unitary representations on it. By "interesting," we meant that these observables transformed in an interesting way under conjugation by the unitary representation. For example, the angular momentum observables rotated amongst each other under conjugation by the unitary representation of $SO(3)$. These observables also generated the unitary representations from which they were derived.

In classical mechanics, we don't have "observables." We also don't have "representations," because phase space isn't a vector space. (Well, it is, but that's not what's important about it. It's certainly not a complex vector space, at any rate.) In place of observables, we have functions on phase space. In place of representations, we have group actions on phase space. Spiritually, it seems as though we should try our old approach: take our group action, look its Lie algebra, and see if there are any interesting ~~observables~~ functions on phase space associated with it.

In order to associate the Lie algebras of group elements with interesting quantities, we will have to make a pretty conceptual observation.

Let's examine one of the few interesting group actions on our phase space $\mathbb{R}^2$, that of $SO(2)$. The action of this group on our phase space is just

$$\begin{pmatrix} q \\ p \end{pmatrix} \to \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$$

which can be rewritten as

$$\begin{pmatrix} q \\ p \end{pmatrix} \to e^{tX} \begin{pmatrix} q \\ p \end{pmatrix}$$

where

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathfrak{so}(3).$$

Here is the conceptual observation: the infinitesimal action of a group on phase space can be pictured using a vector-field.

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{tX} \begin{pmatrix} q \\ p \end{pmatrix} \Bigg|_{t=0} = X \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -p \\ q \end{pmatrix}$$

However, as we saw earlier, a function f on phase space *also* make a vector-field, namely the vector field that sweeps out trajectories along which f is conserved:

$$\frac{\mathrm{d}}{\mathrm{d}t} q = \{q, f\} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t} p = \{p, f\}.$$

The coup is to find the function f that generates the same vector field as X. The answer is

$$f = \frac{1}{2} q^2 + \frac{1}{2} p^2$$

which can be checked by

$$X \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -p \\ q \end{pmatrix} = \begin{pmatrix} -\{q, f\} \\ -\{p, f\} \end{pmatrix}$$

(where the negative is chosen for future convenience). We can see that the value of f doesn't change as points on phase space are acted on by e^{tX}.

We now have a fascinating correspondence between the *Lie algebra elements* of a group that acts on phase space and *functions* on phase space, made possible by associating the vector fields they each create. The map that takes Lie algebra elements to functions on phase space is called the **moment map** μ:

$$\mu : \begin{matrix} \text{Lie algebra elements of a} \\ \text{group that acts on phase space} \end{matrix} \to \text{functions on phase space}$$

For example, with our aforementioned Lie algebra element $X \in \mathfrak{so}(2)$,

$$\mu_X = \frac{1}{2} q^2 + \frac{1}{2} p^2.$$

More generally, if we have a group action G whose elements act on phase space (not necessarily by matrix multiplication)

$$g \cdot \begin{pmatrix} q \\ p \end{pmatrix}$$

the elements $X \in \mathfrak{g}$ will also act on phase space in some way, and the function received from the moment map, μ_X, will be the function that satisfies

$$X \cdot \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -\{q, \mu_X\} \\ -\{p, \mu_X\} \end{pmatrix}$$

This correspondence is not possible for all group actions. It is only possible for group actions that generate "Hamiltonian vector fields." Later we will better characterize for which groups we can do this, but for now its not necessary.

Here's another simple example. The additive group $\mathbb{R}$ acts on phase space by translating the position coordinate. The element $x \in \mathbb{R}$ acts by

$$x \cdot \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q - x \\ p \end{pmatrix}$$

which is, you should note, not matrix multiplication. It should go without saying that $x \cdot \left(\binom{q_1}{p_1} + \binom{q_2}{p_2} \right) \neq x \cdot \binom{q_1}{p_1} + x \cdot \binom{q_2}{p_2}$. Therefore, the Lie algebra of this group will not exponentiate to group elements (at least not in the usual, multiplicative sense of exponentiation). Even so, we can still find the action of the Lie algebra of $\mathbb{R}$ on phase space by acting infinitesimally on it (and then subtracting by the identity, here 0, and dividing by the infinitesimal quantity). The action of the Lie algebra element that generates x is therefore

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\epsilon \cdot \begin{pmatrix} q \\ p \end{pmatrix} - 0 \cdot \begin{pmatrix} q \\ p \end{pmatrix} \right) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\begin{pmatrix} q - \epsilon \\ p \end{pmatrix} - \begin{pmatrix} q \\ p \end{pmatrix} \right) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

The vector field created by this Lie algebra element is

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

which means the conserved quantity along trajectories of this Lie algebra element is just

$$p$$

as

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\{q, p\} \\ -\{p, p\} \end{pmatrix}.$$

Therefore, we can see that momentum is conserved when the position coordinate of a point is translated. (Duh.)

Let's look at what happens to phase space in multiple dimensions. In three dimensions, points in our phase space are given by two vectors:

$$(\vec{q}, \vec{p}) \in \mathbb{R}^6.$$

The Poisson bracket, which still generates trajectories that conserve functions on phase space, is now defined as

$$\{f, g\} \equiv \sum_{i=1}^{3} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i}.$$

$R \in SO(3)$ acts on this phase space in a natural way, defined by

$$R \cdot \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} R\vec{q} \\ R\vec{p} \end{pmatrix}.$$

Each rotation matrix R is specified by an angle t and a unit vector $\hat{n}$. Given our old Lie algebra elements $L_j \in \mathfrak{so}(3)$, we have

$$R = e^{t\hat{n} \cdot \vec{L}}$$

so the infinitesimal action is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{t\hat{n} \cdot \vec{L}} \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix} \bigg|_{t=0} = \begin{pmatrix} (\hat{n} \cdot \vec{L})\vec{q} \\ (\hat{n} \cdot \vec{L})\vec{p} \end{pmatrix}.$$

For simplicity, let's consider the rotation around the z-axis:

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

The appropriate function that produces the same vector field is

$$\mu_{L_z} = p_x q_y - q_x p_y$$

as

$$\{q_x, \mu_{L_z}\} = -q_y \qquad \{q_y, \mu_{L_z}\} = q_x \qquad \{p_x, \mu_{L_z}\} = -p_x \qquad \{p_y, \mu_{L_z}\} = p_x.$$

I hope you agree that μ_{L_z} is (minus) the angular momentum along the z-axis. We can see that a rotation around the z-axis does not change the component of angular momentum in that direction.

All of this has been accomplished without mentioning the Hamiltonian, but now we're ready.

The Hamiltonian h is just a function on phase space that we deem to be important. Time evolution is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} q_i = \{q_i, h\} = \frac{\partial h}{\partial p_i} \qquad \frac{\mathrm{d}}{\mathrm{d}t} p_i = \{p_i, h\} = -\frac{\partial h}{\partial q_i}$$

so we can see that h, A.K.A. energy, is conserved.

The time evolution of any given function f is then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f &= \sum_{i=1}^{3} \frac{\partial f}{\partial q_i} \frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\partial f}{\partial p_i} \frac{\mathrm{d}p_i}{\mathrm{d}t} \\ &= \sum_{i=1}^{3} \frac{\partial f}{\partial q_i} \frac{\partial h}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial h}{\partial q_i} \\ &= \{f, h\}. \end{aligned}$$

The fun begins when we have a Lie group G that leaves h invariant. By "leaves invariant", we mean that for each $g \in G$,

$$h\Big(g \cdot \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix}\Big) = h\Big(\begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix}\Big).$$

Let's say we express g in terms of a Lie algebra element $X \in \mathfrak{g}$,

$$g = e^X$$

(where the exponential is purely formal, and this notation applies even for groups that don't act via matrix multiplication). Using the chain rule, we see

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t}\, h\Big(e^{tX}\cdot\begin{pmatrix}\vec{q}\\ \vec{p}\end{pmatrix}\Big)\Big|_{t=0}\\
&= \sum_{i=1}^{3} -\frac{\partial h}{\partial q_i}\{q_i,\mu_X\} - \frac{\partial h}{\partial p_i}\{p_i,\mu_X\}\\
&= \sum_{i=1}^{3} -\frac{\partial h}{\partial q_i}\frac{\partial \mu_X}{\partial p_i} + \frac{\partial h}{\partial p_i}\frac{\partial \mu_X}{\partial q_i}\\
&= \{\mu_X, h\}.
\end{aligned}$$

This is a very interesting statement, which can be read in two different ways:

1. When points in phase space are moved in such a way that μ_X is constant (such as being acted on by G), h is conserved.

2. When points in phase space are moved such that h is constant (such as evolving in time), μ_X is conserved.

The equivalence of these two statements is often jokingly summed up as

$$\{h,\mu_X\} = 0 = \{\mu_X, h\}.$$

The above statement also constitutes a "proof" that symmetries give conservation laws. Let me flesh it out further:

Recall that acting on phase space by $g = e^X$ is the same as evolving points on phase space according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\vec{q}\\ \vec{p}\end{pmatrix} = \begin{pmatrix}-\{\vec{q},\mu_X\}\\ -\{\vec{p},\mu_X\}\end{pmatrix}.$$

If h doesn't change when points in phase space are acted on by $g = e^X$, we have $\{h,\mu_X\} = 0$. This implies $\{\mu_X, h\} = 0$, which means that μ_X is conserved under time evolution. The triviality of this proof suggests that this is the right language to discuss symmetry and conservation laws.

We can understand the moment map as the equivalent of Noether's theorem in the Hamiltonian picture. Noether's theorem tells us that each infinitesimal group action that leaves the Lagrangian invariant gives us a conserved quantity. The moment map gives us, perhaps, a clearer picture of what symmetries have to do with conservation laws. The moment map approach becomes extra juicy once we quantize. Our group actions will become unitary representations on our Hilbert space (up to a sign), but we'll get to that later.

21 Quantization is a Lie Algebra Representation

We saw that functions on phase space were a Lie algebra with the Poisson bracket acting as the Lie bracket. We also saw that we had a sub Lie algebra when we restricted out functions to polynomials of maximum degree two.

It's no secret that Poisson bracket bears an incredible resemblance to the commutator in classical mechanics. Namely,

$$\{q, p\} = 1$$

looks an awful lot like the canonical commutation relation.

$$[\hat{q}, \hat{p}] = i\hbar\mathbf{1}$$

One might wonder if we can make a Lie algebra representation that maps functions on phase space to operators on a Hilbert space in such a way that the Lie bracket is preserved. The answer is yes, but only up to degree two polynomials.

We will call this Lie algebra representation the "Schrödinger representation" Γ'_S, defined explicitly as

$$\begin{array}{ccc} \Gamma'_S(1) = -i\hbar\mathbf{1} & \Gamma'_S(q) = -i\hat{q} & \Gamma'_S(p) = -i\hat{p} \\ \Gamma'_S(qp) = -i\frac{1}{\hbar}\frac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q}) & \Gamma'_S(q^2) = -i\frac{1}{\hbar}\hat{q}^2 & \Gamma'_S(p^2) = -i\frac{1}{\hbar}\hat{p}^2 \end{array}$$

I'll leave it up to you to check this.

Why can't this be extended to all polynomials? The reason is operator ordering ambiguity. For example, look at how qp gets mapped to $\frac{1}{2}(\hat{q}\hat{p}+\hat{p}\hat{q})$. This is necessary in order to get the correct commutation relations, as $\hat{q}\hat{p} \neq \hat{p}\hat{q}$. (Also note that the combination $\frac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q})$ is self-adjoint like the other operators.) For higher order operators there's no way to do this consistently (although this failure to adhere to the Lie bracket will only occur in higher powers of $\hbar$). The fact that we have to limit ourselves to quadratic operators in order to make a Lie algebra representation is also known as the "Groenewold-van Hove no-go theorem."

There's something else interesting about this Lie algebra representation: up to unitary equivalence, it is unique. This is also true of the sub Lie algebra representation consisting of just q, p, and 1, the Heisenberg representation. The fact that the Heisenberg Lie algebra only has one non-trivial unitary representation is a consequence of the "Stone–von Neumann theorem," and shows that there really is only one way to quantize our system.

This generalizes quite easily to multiple dimensions. For example, in three dimensions, the cross-term polynomials where $j \neq k$ are mapped to

$$\Gamma'_S(q_j q_k) = -i\frac{1}{\hbar}\hat{q}_j\hat{q}_k \qquad \Gamma'_S(p_j p_k) = -i\frac{1}{\hbar}\hat{p}_j\hat{p}_k \qquad \Gamma'_S(q_j p_k) = -i\frac{1}{\hbar}\hat{q}_j\hat{p}_k.$$

22 Quantizing a Group Action

Here is where the moment map really shines.

Lets say we have a group G that acts on the phase space $\mathbb{R}^{2N}$. Say group element $g \in G$ can be expressed as the exponential of a Lie algebra element $X \in \mathfrak{g}$.

$$g = e^X$$

The moment map quantity μ_X is the function on phase space that is conserved under the action of g. That is, g acts on a point in phase space infinitesimally as

$$X \cdot \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix} = \frac{\mathrm{d}}{\mathrm{d}t}\, e^{tX} \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix}\Bigg|_{t=0} = \begin{pmatrix} -\{\vec{q}, \mu_X\} \\ -\{\vec{p}, \mu_X\} \end{pmatrix}.$$

Aside from acting on points in phase space, g also acts on the coordinate *functions* q_i and p_i. The Lie algebra element acts as

$$X \cdot q_i = -\{q_i, \mu_X\} \qquad\qquad X \cdot p_i = -\{p_i, \mu_X\}.$$

Quantization will then take our action of $\mathfrak{g}$ on these coordinate functions to an action of $\mathfrak{g}$ on observables. (Of course, μ_X must be at most a degree 2 polynomial in order to be quantized consistently.)

$$X \cdot \Gamma'_S(q_i) = -[\Gamma'_S(q_i), \Gamma'_S(\mu_X)] \qquad\qquad X \cdot \Gamma'_S(p_i) = -[\Gamma'_S(p_i), \Gamma'_S(\mu_X)]$$
$$X \cdot \hat{q}_i = -[\hat{q}_i, \Gamma'_S(\mu_X)] \qquad\qquad X \cdot \hat{p}_i = -[\hat{p}_i, \Gamma'_S(\mu_X)]$$

Let's now check that this action of $\mathfrak{g}$ on observables is a Lie algebra representation. It's clear that for any operators that are linear combinations of $\hat{q}_i$'s and $\hat{p}_i$'s,

$$\hat{O} = \sum_{i=1}^{N} a_i \hat{q}_i + \sum_{i=1}^{N} b_i \hat{p}_i$$

the action of X is linear in both respects:

$$X \cdot (\hat{O}_1 + \hat{O}_2) = X \cdot \hat{O}_1 + X \cdot \hat{O}_2$$
$$(X + Y) \cdot \hat{O} = X \cdot \hat{O} + Y \cdot \hat{O}.$$

Furthermore, the Jacobi identity ensures that $X \cdot Y - Y \cdot X$ acts in the same way as $[X, Y]$, implying that this really is a Lie algebra representation.

$$\begin{aligned}
X \cdot Y \cdot \hat{O} - Y \cdot X \cdot \hat{O} &= -[Y \cdot \hat{O}, \Gamma'_S(\mu_X)] + [X \cdot \hat{O}, \Gamma'_S(\mu_Y)] \\
&= [[\hat{O}, \Gamma'_S(\mu_Y)], \Gamma'_S(\mu_X)] - [[\hat{O}, \Gamma'_S(\mu_X)], \Gamma'_S(\mu_Y)] \\
&= -[\hat{O}, [\Gamma'_S(\mu_X), \Gamma'_S(\mu_Y)]] \qquad \text{(Jacobi identity)} \\
&= [X, Y] \cdot \hat{O} \checkmark
\end{aligned}$$

(Note that the minus sign accompanying the moment map finally became important.) This actually a manifestation of the more general fact that

$$\{\mu_X, \mu_Y\} = \mu_{[X,Y]}$$

which follows from the Jacobi identity of the Poisson bracket.

In any case, we can exponentiate this Lie algebra representation on operators into a Lie *group* representation on operators.

$$X \cdot \hat{O} = -[\hat{O}, \Gamma'_S(\mu_X)] \quad \xrightarrow{\text{exp}} \quad e^X \cdot \hat{O} = e^{\Gamma'_S(\mu_X)} \hat{O} e^{-\Gamma'_S(\mu_X)}$$

This can be checked via differentiating the action of the Lie group an recovering the Lie algebra.

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{tX} \cdot \hat{O} \bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} e^{\Gamma'_S(\mu_X)} \hat{O} e^{-\Gamma'_S(\mu_X)} \bigg|_{t=0} = -[\hat{O}, \Gamma'_S(\mu_X)] \checkmark$$

The action of G on operators will resemble their pre-quantized counterparts. This is just because, by quantizing, we're just putting hats on our operators. For example, the group of spatial translations acts on the coordinate functions on phase space as

$$q_i \mapsto q_i - a_i.$$

The quantized group action acts in the same way:

$$e^{-i\vec{a}\cdot\hat{p}}\hat{q}_i e^{i\vec{a}\cdot\hat{p}} = \hat{q}_i - a_i$$

(The above equation can be explicitly checked either by using Baker-Campbell-Hausdorff or by acting on a state $\psi(q)$.)

Our group action of G on phase space survives this circuitous process unharmed! After quantization, it becomes an action on operators, where the action is conjugation by the exponentiated moment map element

$$e^{\Gamma'_S(\mu_X)}.$$

Because our group action is perfectly replicated on operators we have the following group homomorphism property.

$$e^X e^Y = e^Z \qquad \Longrightarrow \qquad e^{\Gamma'_S(\mu_X)}e^{\Gamma'_S(\mu_Y)}\hat{O}e^{-\Gamma'_S(\mu_Y)}e^{-\Gamma'_S(\mu_X)} = e^{\Gamma'_S(\mu_Z)}\hat{O}e^{-\Gamma'_S(\mu_Z)}$$

Something else to note is that the observable $i\Gamma'_S(\mu_X)$ will be unchanged, or "conserved," under this group action!

$$e^{\Gamma'_S(\mu_X)}i\Gamma'_S(\mu_X)e^{-\Gamma'_S(\mu_X)} = i\Gamma'_S(\mu_X)$$

Next question: what happens when, on classical phase space, G was a symmetry of the Hamiltonian? We saw that this implied that, for each $X \in \mathfrak{g}$, μ_X was conserved under time evolution.

$$\{h, \mu_X\} = 0$$

Hitting the above expression with the the Schrödinger representation, we find that

$$[\Gamma'_S(h), \Gamma'_S(\mu_X)] = 0 \Longrightarrow [\hat{H}, i\Gamma'_S(\mu_X)] = 0.$$

Therefore, $i\Gamma'_S(\mu_X)$ is also a conserved quantity of the quantum system. (Furthermore, $[\hat{H}, e^{\Gamma'_S(\mu_X)}] = 0$.)

But wait! What if h is not a degree 2 polynomial? Can $\Gamma'_S(h)$ be defined successfully? What about operator ordering ambiguities?

Indeed, there are many Hamiltonians that are not degree 2 polynomials. For example, a very popular Hamiltonian is

$$h = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + V\big((q_1^2 + q_2^2 + q_3^2)^{\frac{1}{2}}\big)$$

where V is some arbitrary function.

Even though the Hamiltonian above has a V term which isn't a degree 2 polynomial, it can still be quantized successfully because there are no p's in the term, ensuring there are no operator ordering ambiguities.

Here is another way to convince yourself that $\{h, \mu_X\} = 0 \implies [\hat{H}, i\Gamma'_S(\mu_X)] = 0$ if h is of the above form. Imagine expanding out V as a Taylor series. The product rule of differentiation yields the following Poisson bracket identity:

$$\{f_1 f_2, g\} = f_1\{f_2, g\} + \{f_2, g\} f_2.$$

The commutator also enjoys a similar identity:

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$$

You can imagine expanding out V as a string of q_i's, and using the Poisson bracket identity to grind down $\{V, \mu_X\}$ until the only brackets left are a bunch of $\{q_i, \mu_X\}$ with some q's out front. You could do the same thing for $[\hat{V}, i\Gamma'_S(\mu_X)]$. As all the $\hat{q}_i$'s commute with each other, $\{V, \mu_X\} = 0$ implies that $[\hat{V}, i\Gamma'_S(\mu_X)] = 0$ because one will just end up as a combination of $\{q_i, p_j\} = \delta_{ij}$'s that will all cancel out, and the other will just end up as a combination of $[\hat{q}_i, \hat{p}_i] = i\hbar\delta_{ij}$ that will all cancel out.

Phew!

Let's take stock. We've found a way to translate group actions of classical systems and their associated conserved quantities into group actions of quantum observable and their associated conserved quantities. We've therefore shown that it is enough to inspect the symmetries and conserved quantities of a classical system in order to find the analogous symmetries and conserved quantities under quantization. We were only able to do this using the conserved quantities given by the moment map.

We can see that symmetries and conserved quantities are a two-way street: symmetries give conserved quantities, and conserved quantities generate symmetries.

I should stop here, but I'll ask a provocative question instead. As $\Gamma'_S(\mu_X)$ is a skew-adjoint matrix (from the definition of Γ'_S) the operator $e^{\Gamma'_S(\mu_X)}$ is unitary. We've already seen that our group action on operators (which acts via conjugation) respects the group homomorphism property. But is the same true of $e^{\Gamma'_S(X)}$? Does the following equation hold?

$$e^X e^Y = e^Z \quad \implies \quad e^{\Gamma'_S(\mu_X)} e^{\Gamma'_S(\mu_Y)} = e^{\Gamma'_S(\mu_Z)}?$$

In other words, does G act as a unitary representation on our Hilbert space?

Much of the time, the answer is 'yes.' For example, in the case of spatial rotations, our quantized moment map quantities will be the orbital angular momentum observables. As we saw, once exponentiated, the angular momentum observables acted as rotations of wave-functions. We can see that, at least in this case, our action of $SO(3)$ on phase space because a unitary representation of $SO(3)$ on the quantum Hilbert space.

But in general, the answer is strangely 'no.' It turns out that this homomorphism property exists only up to a sign that cannot always be removed.

$$e^X e^Y = e^Z \quad \implies \quad e^{\Gamma'_S(\mu_X)} e^{\Gamma'_S(\mu_Y)} = \pm e^{\Gamma'_S(\mu_Z)}$$

The reason for this sign has to do with a double cover lurking under the surface of everything we're doing. (The relevant double cover is the "metaplectic group," which is a double cover of the "symplectic group.") We've actually seen a tentacle of this malicious and mysterious double-cover already. When we showed that the action of the harmonic oscillator Hamiltonian generated a representation of $U(1)$, we were only able to prove this up to a phase. We removed the phase by shifting the Hamiltonian so the ground state energy was 0. So, strangely, the ground state energy of the Harmonic oscillator of $\frac{1}{2}\hbar\omega$ is actually a manifestation of representation theory and topology, albeit in a very mysterious way.

23 Symplectomorphisms and Degree 2 Polynomials

The last section was, in my mind, the culminating section on the moment map. If you understand the last section, you understand the moment map. Having said that, there are a few mysterious things that I would like to briefly clear up for those interested.

We've mentioned that not every group action on phase space has an associated moment map function. Let's give an example of this. For now we'll work in 1 dimension. Consider the multiplicative group $\mathbb{R}^+$, acting on phase space via dilation.

$$(q, p) \to (\lambda q, \lambda p)$$

The vector field created by the infinitesimal action of this group is

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q \\ p \end{pmatrix}$$

but there is simply no function μ for which both of the following equations hold.

$$\frac{\partial \mu}{\partial p} = q \qquad\qquad -\frac{\partial \mu}{\partial q} = p$$

This is because this would imply

$$\frac{\partial}{\partial q}\frac{\partial \mu}{\partial q} = 1 \quad \text{but} \quad \frac{\partial}{\partial p}\frac{\partial \mu}{\partial p} = -1.$$

So why are there some group actions that have a corresponding moment map function, and some that don't?

Let's shrink our question a bit, and only consider the group actions that act via a matrix multiplication and a translation. Such a transformation is called an "affine" transformation.

$$\begin{pmatrix} q \\ p \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} q \\ p \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix}$$

These group actions can be written as 3×3 matrices.

$$\begin{pmatrix} a & b & x \\ c & d & y \\ 0 & 0 & 1 \end{pmatrix}$$

The composition of two group elements is given by matrix multiplication. This is because

$$\begin{pmatrix} A & \vec{v}_1 \\ \mathbf{0} & 1 \end{pmatrix}\begin{pmatrix} B & \vec{v}_2 \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} AB & (A\vec{v}_2 + \vec{v}_1) \\ \mathbf{0} & 1 \end{pmatrix}.$$

We can of course generalize the notion of an affine transformation to N dimensions. These are the transformations that send

$$\begin{pmatrix} \vec{q}_1 \\ \vec{p}_1 \end{pmatrix} \longmapsto \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \vec{q}_1 \\ \vec{p}_1 \end{pmatrix} + \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix}$$

where A, B, C, D are $N \times N$ matrices.

Which affine transformations have associated moment map elements? It turns out that the set of affine transformations that admit a moment map are "symplectomorphisms."

What does that mean?

A "symplectic form" on a phase space $\mathbb{R}^{2N}$ is a map from pairs points on phase space to real numbers.

$$\langle \cdot, \cdot \rangle : \mathbb{R}^{2N} \times \mathbb{R}^{2N} \to \mathbb{R}$$

It must also be anti-symmetric, meaning

$$\langle x_1, x_2 \rangle = -\langle x_2, x_1 \rangle.$$

The simplest symplectic form is given by

$$\langle \begin{pmatrix} \vec{q}_1 \\ \vec{p}_1 \end{pmatrix}, \begin{pmatrix} \vec{q}_2 \\ \vec{p}_2 \end{pmatrix} \rangle = \begin{pmatrix} \vec{q}_1^T & \vec{p}_1^T \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \vec{q}_2 \\ \vec{p}_2 \end{pmatrix}.$$

A symplectomorphism is a transformation that preserves this symplectic form. Physicists often call symplectomorphisms "canonical transformations." An affine transformation preserves the symplectic form if the following equation holds for all points in phase space.

$$\begin{aligned} \begin{pmatrix} \vec{q}_1^T & \vec{p}_1^T \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \vec{q}_2 \\ \vec{p}_2 \end{pmatrix} &= \begin{pmatrix} \vec{q}_1^T & \vec{p}_1^T \end{pmatrix} \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \vec{q}_2 \\ \vec{p}_2 \end{pmatrix} \\ &\quad + \begin{pmatrix} \vec{x}^T & \vec{y}^T \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} \end{aligned}$$

The above equation holds if and only if

$$\begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}$$

which, after matrix multiplication, implies

$$\begin{aligned} A^T C &= C^T A & \qquad B^T D &= D^T B \\ A^T D - C^T B &= \mathbf{1} & \qquad B^T C - D^T A &= -\mathbf{1}. \end{aligned}$$

The affine transformations that satisfy the above equations are the symplectomorphisms. In the case of $N = 1$, these conditions reduce to $\det \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = 1$.

We will now look at simplest elements of this group, find the Lie algebra elements that exponentiate to them, and find the associated moment maps.

$$e^X = \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow X = \begin{pmatrix} 0 & 0 & x \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \mu_X = xp$$

$$e^X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \mu_X = -yq$$

$$e^X = \begin{pmatrix} 1 & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow X = \begin{pmatrix} 0 & b & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \mu_X = b\frac{p^2}{2}$$

$$e^X = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow X = \begin{pmatrix} 0 & 0 & 0 \\ c & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \mu_X = c\frac{q^2}{2}$$

$$e^X = \begin{pmatrix} e^a & 0 & 0 \\ 0 & e^{-a} & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow X = \begin{pmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & 0 \end{pmatrix} \longrightarrow \mu_X = aqp$$

We can see that the most general moment map element for a det $= 1$ affine transformation is necessarily a degree 2 polynomial.

$$aqp + b\frac{p^2}{2} + c\frac{q^2}{2} - yq + xp$$

This generalizes pretty easily to the N dimensional case. The Lie algebra of the group of affine symplectomorphisms of $\mathbb{R}^{2N}$ is in bijection with degree 2 polynomials of $2N$ variables, with the Lie bracket being the Poisson bracket.

I said "bijection," but that's actually not exactly true. Adding a constant to a moment map function does not change the vector field it generates. So really, it's only a bijection up to a constant term.

In any case, we can see that two strange facts we've encountered actually fit together quite nicely. Fact 1: not every group action admits a moment map element. Fact 2: only quadratic polynomials can be quantized consistently.

Apparently, only symplectomorphims have moment map elements, but because the moment map elements are quadratic polynomials, they can always be quantized.

We're now in a position to understand that "double cover" I mentioned a few pages ago. Quadratic polynomials on phase space represent a Lie algebra representation of affine symplectomorphisms. When we apply the Schrödinger representation Γ'_S, this Lie algebra of quadratic polynomials becomes the Lie algebra of quadratic operators on our quantum Hilbert space. However, even though the Lie *algebras* are the same, the Lie *groups* are not. In other words, the operators of the form

$$e^{\Gamma'_S(\mu_X)}$$

do not form a group (where X is in the Lie algebra of the symplectic group). They do, however, form a group up to a sign ambiguity. That is, those operators form a *projective* representation of the symplectic group.

We saw the exact same thing with $SO(3)$ earlier. Because $SO(3)$ is not simply connected, it has non-trivial projective representations. All of the projective representations of $SO(3)$ can be regarded as actual representations of its double-cover

$SU(2)$. Likewise, it turns out that the symplectic group is not simply connected. Just like $SO(3)$, it has non-contractible loops. Its fundamental group is isomorphic to $\mathbb{Z}$ (not $\mathbb{Z}_2$). The double cover of the "symplectic group" is called the "metaplectic group." Global topological information like the fundamental group is not present at the level of the Lie algebra. This is just like how $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic Lie algebras, but $SU(2)$ is the double-cover of $SO(3)$. I won't go into any more detail on the metaplectic group because I don't know that much about it.

In a sense, symplectomorphisms are the natural group of transformations of a classical phase space, and (projective) unitary representations are the natural group of transformations of a quantum Hilbert space. It's very interesting that these two groups of transformations are so intimately connected.