

MATH3001

Quantum Mechanics and Representation Theory

Harry Partridge

Term 1 2021

Contents

1	Overview	2
2	Representations of finite groups	2
2.1	Basic Definitions	2
2.2	Subrepresentations	3
2.3	Equivalent Representations	4
2.4	Examples of Representations	4
2.5	Complete reducibility for finite groups	5
2.6	Schur's lemma	6
2.7	Character Theory	7
3	Representations of Compact Groups	9
3.1	Basic Definitions	9
3.2	Examples of Representations	9
3.3	Maschke's Theorem	9
3.4	Schur's Lemma	9
3.5	Peter Weyl Theorem	9
3.6	Characters	9
4	Lie Theory	9
4.1	Manifolds	9
4.2	Lie groups	10
4.3	Lie Group Representations	11
4.4	Lie Algebras	11
4.5	Lie Algebra Representations	12
4.6	The Lie Algebra of a Lie Group	13
4.7	Covering spaces	13
4.8	The Exponential Map	14
4.9	Lie algebras of $U(n)$, $SL(n, \mathbb{C})$ and $SU(n)$	14
4.10	Irreducible representations of $U(1)$	15
4.11	$\mathfrak{su}(2)$ and $\mathfrak{sl}(2, \mathbb{C})$	15
4.12	Irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$	16
5	Quantum Mechanics	16
5.1	The Quantum Formalism	16
5.2	Measurement	18
5.3	Lie Theory, Representations in Quantum Mechanics	19
5.4	Orbital Angular Momentum	19
5.5	Spin	22
5.6	The periodic table	23
6	Sources	23

1 Overview

The purpose of this report is to explore representation theory and Lie theory, and then examine implications for quantum mechanics. The key result here is that quantum observables arise from representations of Lie groups on a Hilbert space. This process occurs in the following way:

1. Suppose G is a Lie group. The tangent space at the identity $e \in G$ is a vector space that can be made into a lie algebra \mathfrak{g} .
2. Exponentiating elements A of the lie algebra gives lie group elements $g = e^A$ in the connected component of the identity.
3. Suppose G acts unitarily on a Hilbert space \mathcal{H} . Then $e^{(A^\dagger)} = (e^A)^\dagger = e^{-A}$, so $A^\dagger = -A$ and A is skew adjoint.
4. A new self adjoint operator $O = iA$ can be defined, and this will be an observable of the quantum system.

Some examples that we will explore

1. Say \mathcal{H} is $L^2(\mathbb{R}^3)$. Then $SO(3)$ acts on \mathcal{H} by $(\rho(R)f)(x) = f(R^{-1}x)$.
2. Say \mathcal{H} is \mathbb{C}^n . Then what??
3. Say \mathcal{H} is \mathbb{C} . Then $U(1)$ acts on \mathcal{H} by $\rho(e^{i\theta} = e^{ik\theta})$, k an integer.
4. $G = \mathbb{R}$ - time?

2 Representations of finite groups

This section is based primarily on notes by Dexter Chua.

2.1 Basic Definitions

There are several equivalent ways to define a representation of a group. The first is via a homomorphism.

Definition (Representation). For a given group G , a **representation** is a group homomorphism

$$\rho : G \rightarrow GL(V),$$

where V is a vector space over a field \mathbb{F} .

Alternatively, a representation can also be viewed as a linear action on a vector space.

Definition (Linear action). A group G **acts linearly** on a vector space (V, \mathbb{F}) if acts on the set V such that

$$g(\vec{v} + \vec{u}) = g\vec{v} + g\vec{u} \text{ and } g(\lambda\vec{v}) = \lambda g\vec{v}$$

for all $\vec{u}, \vec{v} \in V$ and $\lambda \in \mathbb{F}$. Such an action gives a homomorphism to $GL(V)$ via $\rho(g)(\vec{v}) = g\vec{v}$, and any homomorphism to $GL(V)$ also gives a linear action in an analogous way.

Lastly, a representation can also be viewed as a module over the group algebra $\mathbb{F}G$.

Definition (Group algebra). For a given group G , the group algebra $\mathbb{F}G$ is

$$\mathbb{F}G = \left\{ \sum_{g \in G} \alpha_g g : \alpha_g \in \mathbb{F} \right\}.$$

This is another way of describing a linear action, and gives a homomorphism to $GL(V)$ in the same way.

Regardless of how a representation is defined, it has a **dimension** n which is equal to $\dim_{\mathbb{F}}(V)$. As a choice of basis gives a natural isomorphism from V to \mathbb{F}^n , any representation also has an associated **matrix representation**.

Definition (matrix representation). A matrix representation is a homomorphism R from G to $GL_n(\mathbb{F})$. If V is an n dimensional vector space with $\theta : V \rightarrow \mathbb{F}^n$ an isomorphism, then

$$\rho(g)(\vec{v}) = \theta^{-1}(R(g)\theta(\vec{v}))$$

is an associated representation on V .

If V is a vector space with a representation from G to $GL(V)$, then we say that V is a G -space. Furthermore, if only the identity acts trivially on V (equivalently if the kernel of the homomorphism to $GL(V)$ is $\{e\}$), then we say that the representation is **faithful**.

We now prove some results about the eigenvalues and eigenvectors of representations.

Lemma. Suppose that $\rho : G \rightarrow GL(V)$ is a representation of a finite group G . Then for every $g \in G$, any eigenvalue λ of $\rho(g)$ satisfies $\lambda^n = 1$.

Proof. Suppose that under the associated linear action of G on V , λ is an eigenvalue of g with eigenvector \vec{v} . Then $g^n \vec{v} = \lambda^n \vec{v}$. As G is a finite group, $g^n = e$ for some n , so $\lambda^n = 1$. \square

In particular, $\mathbb{F} = \mathbb{R}$, then all representations have eigenvalues that are either 1 or -1 , and if $\mathbb{F} = \mathbb{C}$, then eigenvalues are roots of unity.

Corollary. Suppose that $\rho : G \rightarrow GL(V)$ is a complex representation of a finite group G . Then for every $g \in G$, $\rho(g)$ is diagonalizable. *does this work in all algebraically closed fields?*

Proof. According the characteristic polynomial of $\rho(g)$ must divide $t^n - 1$. As this polynomial has no repeated roots, $\rho(g)$ must be diagonalizable. \square

2.2 Subrepresentations

As with other algebraic topics, there is a lot that can be learned about representations by decomposing them into ‘simpler’ parts.

Definition (G-subspace). For a given representation $\rho : G \rightarrow GL(V)$ of G , we say that $W \leq V$ is a G -subspace if it is a $\rho(G)$ invariant subspace of V :

$$\rho(g)(W) \leq W$$

for all $g \in G$.

Clearly, any G -subspace gives us another representation.

Definition. If $W \leq V$ is a G -subspace for the representation ρ , then the map $\rho|_W : G \rightarrow GL(V)$ given by $\rho|_W(g) = \rho(g)|_W$ is a **subrepresentation** of ρ .

A G -subspace is **proper** if it is not all of V , and **trivial** if it is just the zero vector.

Definition (Reducible). A representation $\rho : G \rightarrow GL(V)$ is said to be **reducible** if there are proper non-trivial G -subspaces.

A stronger requirement is for V to be the sum of such G -subspaces.

Definition (Decomposable). A representation ρ is said to be **decomposable** if there are proper non-trivial G -subspaces $U, W \leq V$ such that

$$V = U \oplus W.$$

We then say that $\rho = \rho_U \oplus \rho_W$.

A priori, it is not clear that reducibility implies decomposability. This is however, content of Maschke's Theorem, which we will prove later.

2.3 Equivalent Representations

We would like to describe a level of equivalence between two G -spaces. As usual, this is done by considering isomorphisms.

Definition (G-homomorphism). Suppose V and W are two G -spaces with representations ρ_V and ρ_W . If $f : V \rightarrow W$ is a linear map such that

$$f(\rho_V(g)\vec{v}) = \rho_W(g)f(\vec{v}),$$

we say that f is a **G-homomorphism**, and that f **intertwines** ρ_V and ρ_W . If f is also an isomorphism of vector spaces, then we say that V and W are **equivalent** and that f is a **G-isomorphism**.

Evidently, if R_1 and R_2 are two matrix representations of G associated to the same representation ρ on V by isomorphisms θ_1 and θ_2 , then they are equivalent, because $\theta_2\theta_1^{-1}$ is a G -isomorphism that intertwines R_1 and R_2 .

2.4 Examples of Representations

We now list some common representations.

The **trivial representation** 1_G is defined by $1_G(g) = (id : \mathbb{F} \rightarrow \mathbb{F})$ for any field \mathbb{F} . Note that the trivial representation is always one dimensional. One

may also speak of ‘a trivial representation on V ’ which would be taken to mean $\rho(g) = \text{id}_V$.

A more complicated type of representation is a **permutation representation**. For any set X , we can define a vector space $\mathbb{F}X$ as a set of formal sums

$$\mathbb{F}X = \left\{ \sum_{x \in X} \alpha_x x : \alpha_x \in \mathbb{F} \right\},$$

with addition and scalar multiplication defined in the obvious way. If a group G acts on a set X , we can therefore define a linear action on $\mathbb{F}X$ by

$$g \sum_{x \in X} \alpha_x x = \sum_{x \in X} \alpha_x gx.$$

The corresponding representation is called a permutation representation of G .

When $X = G$ with the action of G corresponding to left multiplication, we get a particularly special type of permutation representation ρ_{reg} called the **regular representation**.

Example (Representations of S_3). S_3 acts naturally on the set $X = \{a, b, c\}$, which gives us a three dimensional permutation representation. The natural associated matrix representation are three by three permutation matrices. Contrastingly, the regular representation of S_3 is six dimensional, and its natural associated matrices are a subset of the six by six permutation matrices.

2.5 Complete reducibility for finite groups

In this section, we will show that every finite dimensional representation of a finite group over a field of characteristic zero can be decomposed into a sum of irreducible representations.

Theorem (Maschke’s Theorem). Suppose G is a finite group with a linear action on a finite dimensional vector space V over \mathbb{F} with $(\text{char}(\mathbb{F}), |G|) = 1$. Then if W is a G -subspace of V , there exists a G -subspace U of V such that $V = W \oplus U$.

Proof. Suppose $f : V \rightarrow W$ is an endomorphism on V defined by $f(w) = w$ for $w \in W$. Then because $(\text{char}(\mathbb{F}), |G|) = 1$, $\frac{1}{|G|}$ is well defined, so we can set

$$\bar{f}(v) = \frac{1}{|G|} \sum_{g \in G} gf(g^{-1}v).$$

We claim that $V = W \oplus \ker(\bar{f})$. First of all, since $f(g^{-1}v) \in W$, and W is G invariant, $\text{im}(\bar{f}) \subseteq W$. Furthermore, note that for any $w \in W$, $g^{-1}(w) \in W$, so

$$\begin{aligned} \bar{f}(w) &= \frac{1}{|G|} \sum_{g \in G} gf(g^{-1}w) \\ &= \frac{1}{|G|} \sum_{g \in G} gg^{-1}w \\ &= \frac{1}{|G|} \sum_{g \in G} w \\ &= w. \end{aligned}$$

Therefore $\text{im}(\bar{f}) = W$, and $\bar{f}^2 = \bar{f}$. Then for any $v \in V$, $\bar{f}(\bar{f}(v) - v) = 0$, so $\bar{f}(v) - v \in \ker(\bar{f})$. Hence $v = w + u$ where $u \in \ker(\bar{f})$ and $w = \bar{f}(v) \in W$, so $V = W + \ker(\bar{f})$. Now suppose $y \in W \cap \ker(\bar{f})$. Then $0 = \bar{f}(y) = y$, so $V = W \oplus \ker(\bar{f})$.

Furthermore, if $v \in \ker(\bar{f})$, then for any $h \in G$,

$$\begin{aligned}\bar{f}(hv) &= \frac{1}{|G|} \sum_{g \in G} gf(g^{-1}hv) \\ &= \frac{1}{|G|} \sum_{g' \in G} hg'f(g'^{-1}v) \\ &= h\bar{f}(v) \\ &= 0.\end{aligned}$$

Therefore $\ker(\bar{f})$ is G -invariant, and the theorem is proved. \square

Corollary. Any finite dimensional complex representation of a finite group G is **completely reducible (semisimple)**: there exist irreducible G -subspaces V_1, \dots, V_n such that $V \simeq V_1 \oplus \dots \oplus V_n$.

Proof. If V is one dimensional, it is irreducible so we are done. Otherwise, if V is reducible, use Maschke's theorem to write $V = U \oplus W$, where $\dim(U) < \dim(V)$ and $\dim(W) < \dim(V)$. The theorem then follows by induction on the dimension. \square

2.6 Schur's lemma

Schur's lemma essentially says that irreducible representations of a group cannot be intertwined in a non trivial way. Crucially, the second part of Schur's lemma requires that \mathbb{F} is an algebraically closed field, so it does not apply over \mathbb{R} .

Lemma (Schur's lemma). Suppose V and W are G -invariant vector spaces. Then if $f : V \rightarrow W$ is a G -homomorphism, then f is either the zero map or an isomorphism. In particular, if $W = V$ and \mathbb{F} is an algebraically closed field, then f is a scalar multiple of the identity map.

Proof. Suppose f is a G -homomorphism between G -spaces V and W . Then $\ker(f)$ is a G -subspace of V , so irreducibility implies that either $\ker(f) = V$, in which case $f = 0$, or $\ker(f) = \{0\}$, in which case f is injective. Similarly, if f is not the zero map, then $\text{im}(f) = W$, so f is surjective and therefore an isomorphism.

If $V = W$, then the algebraic closure of \mathbb{F} implies that f has an eigenvalue λ . Therefore $f - \lambda I$ is a G -homomorphism with non zero kernel, and is therefore the zero map. Hence $f - \lambda I = 0$, so $f = \lambda I$. \square

Schur's lemma has a number of important corollaries. For these results, assume that the vector space is over an algebraically closed field.

Corollary. If G is abelian, then its irreducible representations must be one dimensional.

Proof. Suppose $\rho : G \rightarrow GL(V)$ is an irreducible representation of an abelian group G . Then for every $g \in G$, commutativity implies that $\rho(g)\rho(h) = \rho(h)\rho(g)$, so $\rho(h) \in \text{End}(V)$ is a G -homomorphism, so $\rho(h)$ is a scalar multiple of the identity map, and every subspace of V is G -invariant. Therefore V must be one dimensional. \square

Corollary. $Z(G)$ is precisely the set of $g \in G$ such that if ρ is a faithful irreducible representation, then $\rho(g) = \lambda \text{id}_V$, where $\lambda \in \mathbb{F}$.

Proof. If $g \in Z(G)$, then $\rho(g)$ is a G -homomorphism, so according to Schur's lemma $\rho(g) = \lambda \text{id}_V$ for some eigenvalue $\lambda \in \mathbb{F}$.

Conversely, if $\rho(g) = \lambda \text{id}_V$, then $\rho(gh) = \lambda \text{id}_V \rho(h) = \rho(h) \lambda \text{id}_V = \rho(hg)$. As ρ is faithful and therefore also injective, this implies that $gh = hg$, so $g \in Z(G)$. \square

2.7 Character Theory

An essential result from linear algebra says that for square matrices, trace is a similarity invariant: if $R_1 = MR_2M^{-1}$ for some matrix M , then $\text{tr}(R_1) = \text{tr}(R_2)$. Intuitively, this makes sense because the trace is the sum of eigenvalues, and since eigenvalues are defined without reference to a basis, the trace should be basis independent. This means that for a given representation $\rho : G \rightarrow GL(V)$, $\text{tr}(\rho(g))$ is well defined for every $g \in G$. For every representation we therefore have an associated function $\chi : G \rightarrow \mathbb{F}$ given by $\chi = \text{tr} \circ \rho$. This is known as the character of G .

Definition (Character). Suppose $\rho : G \rightarrow GL(V)$ is a representation of a finite group G on a vector space V with field \mathbb{F} . Then we say that ρ affords a **character** $\chi : G \rightarrow \mathbb{F}$ defined by $\chi(g) = \text{tr}(\rho(g))$ for every $g \in G$.

Characters are incredibly useful because they turn out to be a complete invariant: a representation is uniquely determined by its character. We begin with some basic definitions and results.

Definition (Degree). The **degree** of a character $\chi : G \rightarrow GL(V) \rightarrow \mathbb{F}$ is the dimension of the vector space V .

Definition. A character χ is said to be **irreducible** iff its associated representation ρ is irreducible. Similarly, χ is said to be **faithful** iff ρ is.

An initial observation is that if χ has degree n , then for any $g \in \ker(\rho)$, we see that $\chi(g) = \text{tr}(\text{id}_V) = \dim(V) = n$.

For complex representations, we can say a bit more. Thanks to a lemma in section 2.1, we know that any eigenvalue λ of $\rho(g)$ must be a root of unity. This allows us to bound $|\chi(g)|$ above by applying the Cauchy-Schwartz inequality: $|\chi(g)| = |\sum_i \lambda_i| \leq \sum_i |\lambda_i| = \dim(V) = n$. As equality holds when all λ_i are equal, this also says that $|\chi(g)| = n$ iff $\rho(g) = \lambda \text{id}_V$ for an eigenvalue λ .

We can now improve on the final corollary in section 2.6 and say that $Z(G)$ is precisely the set where $|\chi(g)| = n$ for any faithful and irreducible character χ .

Lemma. Suppose ρ and ρ' are two representations with characters χ and χ' . Then if ρ is equivalent to ρ' , then $\chi = \chi'$.

Proof. By equivalence, $\rho(g) = f \circ \rho'(g) \circ f^{-1}$ for some isomorphism f . Hence $\chi(g) = \text{tr}(f \circ \rho'(g) \circ f^{-1}) = \chi'(g)$. \square

Definition (Class functions). For a group G , the space of class functions over \mathbb{F} is the set of functions from G to \mathbb{F} that are constant on the conjugacy classes \mathcal{C}_j of G :

$$\mathcal{C}(G) = \{f : G \rightarrow \mathbb{F} : f(ghg^{-1}) = f(h) \text{ for all } g, h \in G\}.$$

This set is made into a vector space under the normal addition and scalar multiplication of functions. By considering the class functions as column vectors with entries in \mathbb{F} , we can see that $\mathcal{C}(G)$ is isomorphic to \mathbb{F}^k , where k is the number of conjugacy classes of G .

Lemma. Any character χ is a class function.

Proof. By definition, $\chi(ghg^{-1}) = \text{tr}(\rho(g)\rho(h)\rho(g)^{-1}) = \chi(h)$. \square

When taken over the complex numbers, $\mathcal{C}(G)$ can be made into a hermitian inner product space (In fact, $\mathcal{C}(G)$ is then a Hilbert space, as all norms on \mathbb{C}^k are equivalent) by defining

$$\langle f, f' \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{f(g)} f'(g).$$

How can this be extended consistently to all fields where $(\text{char}(\mathbb{F}), |G|) = 1$?

The main result of character theory is that the irreducible characters of a group are orthonormal with respect to this inner product.

Theorem (Completeness of Characters). Over the complex numbers, the irreducible characters of G form an orthonormal basis for $\mathcal{C}(G)$: $\langle \chi_i, \chi_j \rangle = \delta_{ij}$ and every $f \in \mathcal{C}(G)$ can be written as $\sum_i \alpha_i \chi_i$ for some coefficients α_i .

Proof. TO DO. \square

Corollary. The number of conjugacy classes of G is the number of complex irreducible characters.

Proof. The irreducible characters form a basis for $\mathcal{C}(G) \simeq \mathbb{C}^k$, so there must be exactly k such characters. \square

Corollary. Complex representations of finite groups are uniquely determined by their characters.

Proof. TO DO. \square

Corollary (Irreducibility Criterion). A complex representation ρ of a finite group G affording character χ is irreducible iff $\langle \chi, \chi \rangle = 1$.

Proof. Suppose ρ is irreducible. Then so is χ and $\langle \chi, \chi \rangle = 1$.

If $\langle \chi, \chi \rangle = 1$, then since $\chi = \sum m_i \chi_i$, we must have that $\sum m_i^2 = 1$ for some integers m_i . Then exactly one $m_i = 1$, and the rest are zero, so χ is irreducible. \square

Theorem. sum of squares of dimensions is size of G .

Proof. TO DO. □

To conclude this section, we look at an example.

Example (Irreducible representations of S_3). TO DO.

3 Representations of Compact Groups

3.1 Basic Definitions

Definition (Hilbert Space). A vector space endowed with a inner product is a Hilbert space if it is complete under the norm induced by the inner product.

Definition (Unitary group). Suppose \mathcal{H} is a Hilbert space. The unitary group $U(\mathcal{H})$ is the set of all bounded linear operators $\mathcal{H} \rightarrow \mathcal{H}$ that satisfy $\langle Uv, Uw \rangle = \langle v, w \rangle$ for all $v, w \in \mathcal{H}$.

Definition (Unitary representation). Suppose G is a compact group and \mathcal{H} is a Hilbert space. A **strongly continuous unitary representation** of G on \mathcal{H} is a group homomorphism $\pi : G \rightarrow U(\mathcal{H})$ such that

3.2 Examples of Representations

3.3 Maschke's Theorem

3.4 Schur's Lemma

3.5 Peter Weyl Theorem

3.6 Characters

4 Lie Theory

4.1 Manifolds

We begin by reviewing some manifold theory.

Definition (Manifold). Let M be a set with a second countable Hausdorff topology τ . A n dimensional smooth structure on M consists of the following elements:

1. A collection of connected open sets $V_\alpha \in \tau$ whose union is M .
2. A collection of homomorphisms $\phi_\alpha : V_\alpha \rightarrow U_\alpha$, where U_α is a connected open subset of \mathbb{R}^n . These are known as charts

M is made into an n dimensional **manifold** if the charts are compatible: for all ϕ_α and ϕ_β , we must have $\phi_\alpha(V_\alpha \cap V_\beta)$ open in \mathbb{R}^n and

$$\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(V_\alpha \cap V_\beta) \rightarrow \phi_\alpha(V_\alpha \cap V_\beta)$$

must be smooth maps on \mathbb{R}^n .

Definition (Smooth map between manifolds). Suppose that M with charts ϕ_α and N with charts ψ_β are smooth m and n dimensional manifolds respectively. A function $f : M \rightarrow N$ is smooth iff for all ϕ_α and ψ_β , we have that $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ is a smooth map from \mathbb{R}^m to \mathbb{R}^n in the usual sense.

Definition (Diffeomorphism). A smooth map f between two manifolds M and N is a **diffeomorphism** if f^{-1} exists and is also a smooth map.

In particular, a diffeomorphism is a special type of homeomorphism between topological spaces.

We omit the proof of the following proposition.

Proposition (Product manifold). Suppose that M with topology τ_M and smooth structure (V_α, ϕ_α) and N with topology τ_N and smooth structure (U_β, ψ_β) are smooth m and n dimensional manifolds respectively. Then

1. $M \times N$ with the product topology $\tau_M \times \tau_N$ is a manifold when endowed with charts $\phi_{\alpha\beta} : V_\alpha \times U_\beta \rightarrow \mathbb{R}^m \times \mathbb{R}^n : (v, u) \mapsto (\phi(v), \psi(u))$.
2. If X and Y are submanifolds of M and N respectively, then $X \times Y$ is a submanifold of $M \times N$.

In particular, for any $x \in M$, $\{x\} \times N$ is a submanifold of $M \times N$.

4.2 Lie groups

A Lie group is a group G that is also a smooth manifold.

Definition (Lie Group). Let G be a group that is also a manifold under a given topology τ and smooth structure (V_α, ϕ_α) . Then G is a **Lie group** iff

1. $m : G \times G \rightarrow G : (g, h) \mapsto gh$ and
2. $\text{inv} : G \rightarrow G : g \mapsto g^{-1}$

are both smooth maps between manifolds.

Definition (Lie group homomorphism). A map between two lie groups G and H is a **lie group homomorphism** if it is a homomorphism of groups that is also a smooth map.

A homomorphism of Lie groups f is an isomorphism if f^{-1} exists and is also a Lie group homomorphism. Equivalently, an isomorphism of Lie groups is a diffeomorphism that is also a homomorphism in both directions.

For a given n dimensional vector space V , the general linear group $GL(V)$ is a Lie group with manifold structure inherited from \mathbb{R}^{n^2} .

Definition (Lie group action). Let G be a Lie group and let X be a manifold. A **Lie group action** of G on X is a group action $G \times X \rightarrow X : (g, x) \mapsto gx$ that is also a differentiable map between manifolds.

Definition (Left and right actions). For any g an element of a Lie group G , the **left action** of g is $L_g : G \rightarrow G : h \mapsto gh$. Similarly, the **right action** of g is $R_g : G \rightarrow G : h \mapsto hg$.

Theorem. For any element $g \in G$, the left and right actions of G are diffeomorphisms on G .

Proof. The left action of g is the restriction of the multiplication map m to the Lie subgroup $\{g\} \times G = \{(g, h) : h \in G\}$. This can be written as $L_g = m \circ \iota$, so as the composition of smooth maps, L_g is also smooth. $L_{g^{-1}}$ is an obvious smooth inverse, so L_g is a diffeomorphism. The argument for R_g is the same. \square

Definition (Conjugation action). For any $g \in S$, the **Conjugation action** $\Psi_g : G \rightarrow G$ of g is composed of left and right actions: $\Psi_g(h) = L_g \circ R_g^{-1}(h) = ghg^{-1}$.

As a composition of diffeomorphisms, the conjugation action is itself a diffeomorphism. From elementary group theory, we know that the conjugation action is also a homomorphism with an inverse that is also a homomorphism. Hence the conjugation action is in fact a Lie group isomorphism.

4.3 Lie Group Representations

As expected, A Lie group representation is a Lie group homomorphism from G to $GL(V)$ for some vector space V . This can also be viewed as a linear Lie action on a vector space V where V is made into a manifold in the natural way. The conjugation map Ψ_g gives a particularly nice Lie group representation.

Since Ψ_g fixes the identity e in G , the derivative $(d\Psi_g)_e : T_e G \rightarrow T_e G$ is an endomorphism on $T_e G$. Furthermore, because Ψ_g is a Lie group isomorphism, Ψ_g^{-1} is smooth, so $(d\Psi_g)_e^{-1}$ exists. Therefore $(d\Psi_g)_e$ is actually an automorphism on $T_e G$. This motivates the following definition.

Definition (Adjoint representation of a Lie group). Suppose G is a Lie group. The **adjoint representation** $\text{Ad} : G \rightarrow \text{Aut}(T_e G)$ of G is the map $\text{Ad}(g) = (d\Psi_g)_e$.

4.4 Lie Algebras

We begin with the definition of a Lie algebra.

Definition (Lie algebra). A vector space V over a field \mathbb{F} is a Lie algebra if it is endowed with a multiplication $V \times V \rightarrow V$ denoted $(x, y) \mapsto [x, y]$ such that

1. The bracket operation is bilinear.
2. $[x, x] = 0$ for all $x \in V$.
3. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$

The third requirement is known as the Jacobi identity.

The first thing to note is that applying the first and second requirements to $[x + y, x + y]$ we see that $[x, y] = -[y, x]$, so the bracket is anti commutative. We also note that in general the bracket does not have to be associative (the Jacobi identity can be used to show that associativity will hold only when $[a, [b, c]] = 0$) for all a, b, c .

Our first example of a Lie algebra is a very important one.

Definition (General linear algebra). For a given vector space V , the set of all endomorphisms on V form an algebra with scalar multiplication from the field \mathbb{F} in V . This vector space can be made into the **general linear algebra** $\mathfrak{gl}(V)$ by replacing the standard product with a bracket operation where

$$[x, y] = xy - yx$$

for all $x, y \in V$.

Any vector space V can be made into a trivial lie algebra by defining $[x, y] = 0$ for all $x, y \in V$. Such a lie algebra is said to be **abelian**.

Definition (Lie algebra homomorphism). Suppose \mathfrak{g} and \mathfrak{h} are two Lie algebras. A function $\tau : \mathfrak{g} \rightarrow \mathfrak{h}$ is a **Lie algebra homomorphism** if for all $X, Y \in \mathfrak{g}$,

$$\tau([X, Y]) = [\tau(X), \tau(Y)]$$

and

$$\tau(\alpha X + \beta Y) = \alpha \tau(X) + \beta \tau(Y).$$

As usual, a homomorphism is a monomorphism, epimorphism or isomorphism if it is respectively injective, surjective or bijective. Two Lie algebras \mathfrak{g} and \mathfrak{h} are said to be **isomorphic** if there exists a Lie algebra isomorphism between them.

Definition (Subalgebra). Suppose \mathfrak{g} is a Lie algebra. A vector subspace \mathfrak{h} of \mathfrak{g} is a Lie subalgebra if it is closed under the bracket operation in \mathfrak{g} : $[x, y] \in \mathfrak{h}$ for all $x, y \in \mathfrak{h}$.

Definition (ideal). An **ideal** \mathfrak{a} of a Lie algebra \mathfrak{g} is a subalgebra that has total closure: for all $x \in \mathfrak{g}$ and $a \in \mathfrak{a}$, $[a, x] \in \mathfrak{a}$. According to skew symmetry, this is equivalent to requiring that $[x, a] \in \mathfrak{a}$.

When \mathfrak{a} is an ideal of a Lie algebra \mathfrak{g} , the quotient vector space $\mathfrak{g}/\mathfrak{a}$ can be made into a Lie algebra under the bracket $[x + \mathfrak{a}, y + \mathfrak{a}] = [x, y] + \mathfrak{a}$.

4.5 Lie Algebra Representations

to do: expand.

Just as a group representation was a group homomorphism into the general linear group, a lie algebra representation is a lie algebra homomorphism into the general linear algebra.

Definition (Representation of a Lie algebra). Suppose \mathfrak{g} is a Lie algebra. A **representation** of \mathfrak{g} on a vector space V is a Lie algebra homomorphism

$$\tau : \mathfrak{g} \rightarrow \mathfrak{gl}(V).$$

Suppose \mathfrak{g} is a Lie algebra, and $x \in \mathfrak{g}$. The map $\text{ad } x : \mathfrak{g} \rightarrow \mathfrak{g} : y \mapsto [x, y]$ is a derivation on \mathfrak{g} . This gives rise to the following important Lie algebra representation.

Definition (Adjoint representation of a Lie algebra). The map $\text{Ad} : \mathfrak{g} \rightarrow \text{Der}(\mathfrak{g}) : x \mapsto \text{ad } x$ is the **adjoint** representation of \mathfrak{g} .

4.6 The Lie Algebra of a Lie Group

It turns out that every Lie group has an associated Lie algebra. This is very useful, because often Lie algebras are easier to study than Lie groups. The key result is that the adjoint representation of a Lie group allows for the definition of a bracket operation on the tangent space.

Recall that the adjoint representation Ad of a Lie group G is a map from G to the automorphisms on $T_e G$. $\text{Aut}(T_e G)$ is an open subset of $\text{End}(T_e G)$, so the tangent space of $\text{Aut}(T_e G)$ at the identity $I = \text{id}_{T_e G}$ can be naturally identified with $\text{End}(T_e G)$. As such, the derivative of Ad is a map

$$\text{ad} = (d\text{Ad})_I : T_e G \rightarrow \text{End}(T_e G).$$

Theorem. For a Lie group G , the vector space $T_e G$ is made into a Lie algebra \mathfrak{g} by the bilinear map

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g} : [x, y] = \text{ad}(x)(y).$$

Proof. We verify the three axioms for Lie algebras.

1. ρ
2. $[x, x] = (d\text{Ad})_I(x)(x)$.
- 3.

□

There is also a one to one correspondence between homomorphisms between Lie groups and homomorphisms between their associated Lie algebras.

Theorem. Suppose G and H are Lie groups with G simply connected. A homomorphism $\rho : G \rightarrow H$ is uniquely determined by its differential $(d\rho)_e : T_e G \rightarrow T_e H$. Furthermore, a linear map $T_e G \rightarrow T_e H$ is the differential of a homomorphism $\rho : G \rightarrow H$ iff it preserves the bracket operation.

Proof.

□

4.7 Covering spaces

to do: define covering space map. Prove the theorems.

Theorem. Suppose G is a Lie group, H a connected manifold and $\phi : H \rightarrow G$ is a covering space map. If $e' \in \phi^{-1}(\{e\})$, then there is a unique Lie group structure on H such that e' is the identity and ϕ is a homomorphism of Lie groups. In this case, the kernel of ϕ is in the center of H .

Theorem. Suppose H is a Lie group, and $\Gamma \subseteq Z(H)$ is a discrete subgroup of its center. Then there is a unique Lie group structure on $G = H/\Gamma$ such that the quotient map $H \rightarrow G$ is a Lie group homomorphism.

Theorem. Suppose G is a connected Lie group, and $U \subseteq G$ is any neighborhood of the identity. Then U generates G .

Proof. First of all, note that any neighborhood of a point contains an open neighborhood of that point, so if necessary we replace U with open subset. U^{-1} is also an open neighborhood of the identity, so if necessary, we further replace U with $U \cap U^{-1}$ such that $U = U^{-1}$.

Let S be the set generated by U . For any $g \in S$, the left action of g is a diffeomorphism, so gU is open because U is. Now write $H = G \setminus S$. If $hu \in S$ for any $u \in U$, $h \in H$, then $huu^{-1} = h$ is also in S , so we must have that hU is disjoint with S . Hence as hU is also open, $H' = \bigcup_{h \in H} hU$ is a open subset of G that is disjoint from S with $G = H' \cup S$. By the connectedness of G , since S is nonempty, we must have $G = S$. \square

4.8 The Exponential Map

We now illustrate how a Lie group homomorphism induces a Lie algebra homomorphism via the exponential map.

4.9 Lie algebras of $U(n)$, $SL(n, \mathbb{C})$ and $SU(n)$

Definition ($U(n)$). $U(n)$ is the group of unitary matrices in n dimensions:

$$U(n) = \{X \in GL(n, \mathbb{C}) : XX^\dagger = 1\}.$$

If we write $\mathfrak{u}(n)$ for the Lie algebra of $U(n)$, then for all $X \in \mathfrak{u}(n)$ and $t \in \mathbb{R}$, $e^{tX} \in U(n)$, so $e^{tX}e^{tX^\dagger} = 1$. Differentiating this expression with respect to t yields the condition $X + X^\dagger = 0$. Therefore, elements of $\mathfrak{u}(n)$ must be skew adjoint matrices. This inspires the following proposition.

Proposition. The Lie algebra $\mathfrak{u}(n)$ of $U(n)$ is precisely the set of skew adjoint $n \times n$ complex matrices.

Proof. For a skew adjoint complex matrix, diagonal entries are imaginary and entries in the upper triangular section are completely determined by the lower triangular section. Therefore, there are $n + 2\frac{n^2-n}{2} = n^2$ real parameters in such a matrix.

On the other hand, because the tangent space has the same dimension as the manifold, the dimension of $\mathfrak{u}(n)$ must be the same as $U(n)$. Unitary matrices are embedded in $2n^2$ real dimensional space, but are subject to n real constraints normalizing the columns, and $\frac{n^2-n}{2}$ complex constraints requiring that columns are orthogonal. Therefore $\mathfrak{u}(n)$ has dimension $2n^2 - n - 2\frac{n^2-n}{2} = n^2$ as well.

Therefore, every skew adjoint matrix must also be in $\mathfrak{u}(n)$ because $\mathfrak{u}(n)$ is a subset of the skew adjoint matrices with the same dimension. \square

Definition ($SL(n, \mathbb{C})$). $SL(n, \mathbb{C})$ is the group of $n \times n$ complex matrices with determinant 1:

$$SL(n, \mathbb{C}) = \{X \in GL(n, \mathbb{C}) : \det(X) = 1\}.$$

Once again, we can use the exponential map to examine the Lie algebra: for any element $X \in \mathfrak{sl}(n, \mathbb{C})$, we have $\det(e^X) = 1$, so as $\det(e^X) = e^{\text{tr}(X)}$, we have that $\text{tr}(X) = 0$.

Proposition. The Lie algebra $\mathfrak{sl}(n)$ of $SL(n, \mathbb{C})$ is precisely the set of traceless $n \times n$ complex matrices.

Proof. As both the trace and determinant map a matrix to the complex numbers, both conditions impose two constraints on the matrices.

Therefore, every traceless $n \times n$ complex matrix must also be in $\mathfrak{sl}(n)$ because $\mathfrak{sl}(n)$ is a subset with the same dimension. \square

Definition ($SU(n)$). $SU(n)$ is the intersection of $SL(n, \mathbb{C})$ and $U(n)$:

$$SU(n) = \{X \in GL(n, \mathbb{C}) : XX^\dagger = 1, \det(X) = 1\}.$$

As such, the lie algebra $\mathfrak{su}(n)$ of $SU(n)$ is just the intersection of $\mathfrak{sl}(n, \mathbb{C})$ and $\mathfrak{u}(n)$.

Proposition. $SL(n, \mathbb{C})$ is the set of $n \times n$ complex skew adjoint traceless matrices:

$$\mathfrak{su}(n) = \{X \in M(n, \mathbb{C}) : X^\dagger = -X, \text{tr}(X) = 0\}.$$

4.10 Irreducible representations of $U(1)$

$U(1)$ is the set of unitary one by one matrices. Therefore $U(1)$ is just the unit circle in the complex numbers. This group is clearly abelian, so by the first corollary of Schur's lemma, the irreducible representations of $U(1)$ must be one dimensional. This implies that any irreducible representation ρ of $U(1)$ is actually just an automorphism on $U(1)$. This gives us our first lemma.

Lemma. Any irreducible representation ρ of $U(1)$ is of the form

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

Proof. As argued above, we must have that ρ \square

4.11 $\mathfrak{su}(2)$ and $\mathfrak{sl}(2, \mathbb{C})$

Recall that $\mathfrak{su}(2)$ is the set of skew symmetric traceless 2×2 complex matrices:

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

On the other hand, since $\mathfrak{sl}(2, \mathbb{C})$ is just the set of traceless 2×2 complex matrices, we see that

$$\begin{aligned} \mathfrak{sl}(2) &= \begin{pmatrix} x - iy & a + ib \\ c + id & -x + iy \end{pmatrix} \\ &= \begin{pmatrix} -iy & -s + iu \\ s + iu & iy \end{pmatrix} + i \begin{pmatrix} -ix & -v + it \\ v + it & ix \end{pmatrix} \end{aligned}$$

where $s = \frac{1}{2}(c - a)$, $t = -\frac{1}{2}(c + a)$, $u = \frac{1}{2}(d - b)$ and $v = \frac{1}{2}(d + b)$. Hence, it is not hard to see that $\mathfrak{sl}(2, \mathbb{C}) = \mathfrak{su}(2) \oplus i\mathfrak{su}(2)$. We will therefore work primarily with $\mathfrak{sl}(2, \mathbb{C})$, and derive representations of $\mathfrak{su}(2)$ as a consequence.

A basis for $\mathfrak{sl}(2, \mathbb{C})$ is given by

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Recalling that the bracket is defined by $[x, y] = xy - yx$, we see that these matrices obey the commutation relations

$$[h, e] = -2e, \quad [h, f] = 2f, \quad [e, f] = -h.$$

4.12 Irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$

Suppose that $\mathfrak{sl}(2, \mathbb{C})$ acts linearly on a finite dimensional vector space V . We define the *weight space* $V_\lambda = \{v \in V : hv = \lambda v\}$ for a given *weight* λ . Clearly, if λ is not an eigenvalue of h , then $V_\lambda = \{0\}$.

Proposition. For any $v \in V_\lambda$, $ev \in V_{\lambda-2}$ and $fv \in V_{\lambda+2}$.

Proof. According to the commutation relations given in the above section, $he - eh = -2e$, so $h(ev) = (eh - 2e)v = (\lambda - 2)ev$. Similarly, $hf - fh = 2f$, so $h(fv) = (fh + 2f)v = (\lambda + 2)fv$. \square

The point of this proposition is to show that these weight spaces of h form chains with weights incrementing by 2. The next two propositions shows that there is exactly one such chain.

Proposition. Suppose that $\mathfrak{sl}(2, \mathbb{C})$ acts linearly on a finite dimensional vector space V . Then there exists an eigenvector $w \in V$ for H such that $ew = 0$.

Proof. As \mathbb{C} is algebraically closed, h must have at least one eigenvalue with an eigenvector v . Then if v, ev, e^2v, \dots are all non zero, then they must all be eigenvectors of h with distinct eigenvalues. This would be an infinite sequence of independent vectors in V , which is impossible because V is finite dimensional. Hence there exists an eigenvector $w = e^k v$ for h such that $ew = e^{k+1}v = 0$ for some k . \square

If w is an eigenvector of h with eigenvalue λ such that $ew = 0$, then we say that w is a maximal vector of V and that λ is a maximal weight.

Proposition. If w is a maximal vector of a vector space V with dimension n , then $w, fw, \dots, f^{n-1}w$ is a basis for V .

Proof. \square

5 Quantum Mechanics

5.1 The Quantum Formalism

Classically, a physical system consisted of the following components:

1. A set S of possible states. This set was a finite dimensional real vector space, with each coordinate corresponding to the value of a particular quantity. The original example of this is description of a single particle in Newtonian mechanics, in which S is a two dimensional real vector space, with the first dimension corresponding position and the second to velocity.

2. A set Ω of observable quantities. These quantities were expressed as a function from S to \mathbb{R} . For example, if a particle of mass m was modeled in a Newtonian system, then the energy of a particle would be given as $E : S \rightarrow \mathbb{R} : E(x, v) = mv^2/2$.
3. A rule which describes how the state evolves through time. For example, if a particle with mass m under the influence of a constant force F was modeled by Newtonian mechanics, then $(\dot{x}, \dot{v}) = (v, \frac{F}{m})$.

Embedded in component 2 is the assumption that every observable must always be completely described by a single real value. Unfortunately however, several experiments in the early 20th century showed this to be insufficient when modeling systems which are precise enough to deal with individual particles. As a result, the founders of quantum mechanics developed a system where

1. S is replaced with a Hilbert space \mathcal{H} . Examples include taking $\mathcal{H} = L^2(\mathbb{R}^3)$ to model a spin-less particle in three dimensions, $\mathcal{H} = \mathbb{C}^2$ for a two state system like the spin of an electron, $\mathcal{H} = L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$ for two particles in one dimension, and $\mathcal{H} = L^2(\mathbb{R}) \otimes \mathbb{C}^3$ for a particle in one dimension with three possible spin states.
2. Observable quantities become hermitian operators on \mathcal{H} . To see what this means in a physical context, note that hermitian operators are diagonalizable. Therefore, specifying a hermitian operator is equivalent to specifying a set of real eigenvalues (the possible ‘well defined’ values of the observable) and an orthonormal basis for \mathcal{H} (the states corresponding to those eigenvalues). The benefit of this description is that it allows for the state of a particle to have nonzero components in multiple eigenspaces - i.e. the particle can be in a ‘superposition’ of eigenstates.
3. The equation defining evolution of the state through time is replaced with the Schrödinger equation:

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

where H is a distinguished observable called the Hamiltonian, and \hbar is plank’s constant divided by 2π .

An important consequence of 2 is that if a system is in state $|\phi\rangle \in \mathcal{H}$, then all measurable quantities will be unchanged if the state changes to $\lambda |\phi\rangle$ for any scalar $\lambda \neq 0$ (Note that $|0\rangle$ does not correspond to a physical state, because it has no components in any eigenspace of any observable). As such, a state is actually better associated with a vector in \mathcal{H}/\mathbb{C} .

When the eigenvalue spectrum of an operator is discrete, we find that the corresponding measurable becomes ‘quantized’, and cannot take on the full range of values we would expect from classical mechanics. This is however, not always the case - it is common for observables such as position X and momentum P to have a continuous spectrum of eigenvalues. Unfortunately there is considerably more technical detail required for a precise mathematical treatment of this situation. For example, a continuous spectrum of eigenvalues would require that there is also an continuous set of orthonormal eigenvectors,

which is not possible in typical Hilbert spaces such as \mathbb{C}^n or $L^2(\mathbb{R}^3)$. The physics approach usually handles this case by pretending that the Hilbert space also contains an object known as the Dirac delta function δ , which is zero everywhere except for at point where it is infinite. A more mathematical approach requires the notion of a rigged Hilbert space, which we do not discuss.

Before proceeding further, we consider an easy example. Suppose that \mathcal{H} is two dimensional and H has eigenvectors $|\psi_1\rangle$ and $|\psi_2\rangle$, with eigenvalues E_1 and E_2 respectively. Our aim is to find $|\psi(t)\rangle$ for any t . We start by writing $|\psi(t)\rangle = \alpha(t)|\psi_1\rangle + \beta(t)|\psi_2\rangle$, for arbitrary functions α, β . Applying the Schrödinger equation,

$$i\hbar\alpha'(t)|\psi_1\rangle + i\hbar\beta'(t)|\psi_2\rangle = E_1\alpha(t)|\psi_1\rangle + E_2\beta(t)|\psi_2\rangle.$$

Equating components, we see that $\alpha(t) = Ae^{-iE_1t/\hbar}$ and $\beta(t) = Be^{-iE_2t/\hbar}$ for constants A, B determined by the initial conditions. The solution is therefore $|\psi(t)\rangle = Ae^{-iE_1t/\hbar}|\psi_1\rangle + Be^{-iE_2t/\hbar}|\psi_2\rangle$.

This example highlights an important consequence of the linearity of the Schrödinger equation: eigenstates of H evolve independently through time. It turns out that this simple observation is the source of significant philosophical debate, which we discuss briefly in the next section.

5.2 Measurement

The formalism of quantum mechanics becomes useful when it can be applied to deduce results of experiments that can be conducted by humans. The problem is that any measurement system (which necessarily includes the people doing the measuring) has at least $\sim 10^{27}$ atoms. Unfortunately, it is impractical to model the evolution of such a system with Schrödinger's equation. Physicists wanting to make useful predictions therefore imagine that when a quantum system interacts with a 'measuring device', the system collapses into a well defined state. This is known as the Born rule.

Principle (Born Rule). Suppose a particle is in the state $|\psi\rangle = \sum \alpha_\omega |\omega\rangle$ where $|\omega\rangle$ are the eigenvectors of an observable Ω . Measurement of Ω will yield the value ω with probability

$$P(\omega) = \frac{|\langle\omega, \psi\rangle|^2}{\langle\psi, \psi\rangle} = \frac{|\alpha_\omega|^2}{\sum |\alpha_\omega|^2}.$$

This measurement will cause the state of the particle to change from $|\psi\rangle$ to $|\omega\rangle$.

Another perspective says that when a measurement is performed, the quantum system must simply be expanded to include the state of the measuring device. For example, if an experiment to determine the outcome of a 'quantum' coin flip was conducted, the resulting state would be

$$\alpha|H\rangle \otimes |H \text{ seen}\rangle + \beta|H\rangle \otimes |T \text{ seen}\rangle + \gamma|T\rangle \otimes |H \text{ seen}\rangle + \delta|T\rangle \otimes |T \text{ seen}\rangle$$

for some constants $\alpha, \beta, \gamma, \delta$ (β and γ are probably very small). In reality, the dimensionality of the new system is actually very high. As a result, each of these four 'macroscopic' states will be very close to an eigenstate of the hamiltonian

H . As eigenstates of H evolve independently from each other, we therefore see that each of these four states are effectively isolated from each other. This process is studied more rigorously in the subject of decoherence, which has been an active area of research since the 1980's.

Under this perspective, there was no 'collapse' - all of the outcomes of the experiment did in fact occur, but they are no longer interacting. This has a number of troubling consequences, most notably that there are an unimaginably number of slightly different 'versions' of reality.

It can therefore be argued that the notion of collapse (the Copenhagen interpretation) is 'simpler' because it doesn't require the existence of infinite noninteracting states which we can never detect. On the other hand it can also be argued that 'no collapse' (the many worlds interpretation) is simpler because it doesn't require a poorly defined, non-deterministic process to be a part of the theory. Clearly, this becomes a matter of philosophy.

5.3 Lie Theory, Representations in Quantum Mechanics

It turns out that for a quantum system modeled by a Hilbert space \mathcal{H} , most of the interesting observables arise from the symmetries of \mathcal{H} . In this section, we give an outline of how this works, and the remaining sections examine several examples.

Suppose that a compact group G has a strongly continuous unitary representation $\pi : G \rightarrow U(\mathcal{H})$ for some Hilbert space \mathcal{H} . We say that G is a group of symmetries for \mathcal{H} because it preserves the inner product (and therefore also the probabilities of various physical measurements). According to the Peter Weyl theorem, \mathcal{H} decomposes into a direct sum of finite dimensional unitary representations of G .

In the section on Lie theory, we showed that the Lie algebra $\mathfrak{u}(n)$ of $U(n)$ is the set of $n \times n$ complex skew adjoint matrices. Furthermore, we showed that any Lie group representation π has an associated Lie algebra representation π' . Therefore, for any $X \in \mathfrak{g}$, $\pi'(X)$ is a skew adjoint, and acts on \mathcal{H} , so $i\pi'(X)$ is a hermitian operator (observable) on \mathcal{H} .

These observable transform in a particularly nice way under the action of $\pi(G)$ on \mathcal{H} . $\pi(g)\pi'(X)\pi(g)^{-1} = \pi'(gXg^{-1})$.

To summarize, the action of a Lie group on a Hilbert space gives rise to a set of observables that respect the symmetries encapsulated by that Lie group. As we will see in the following sections, it just so happens that the observables generated in this manner turn out to be the observables of most interest to us.

5.4 Orbital Angular Momentum

Suppose that we have a Hilbert space \mathcal{H} that is modeling a system in three dimensional space. Then the natural unitary representation of $SO(3)$ on \mathcal{H} will generate the angular momentum operators.

To illustrate this process, we are going to study a single spin-less particle, which we will model by choosing \mathcal{H} to be $L^2(\mathbb{R}^3, \mathbb{C})$. In this situation, we imagine a state $\phi \in \mathcal{H}$ represents a probability cloud where $|\phi(\vec{x})|^2$ is the probability of being found around the point \vec{x} . Then it isn't too hard to see

that the representation $\pi : SO(3) \rightarrow GL(H)$ defined by

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

is ‘natural’ in the sense that it corresponds to rotating the probability cloud ϕ by R . To see that this is also a unitary representation, note that for any $R \in SO(3)$, Jacobian is 1 so

$$\langle \pi(R)\psi, \pi(R)\phi \rangle = \int \psi(R^{-1}\vec{x})^* \phi(R^{-1}\vec{x}) d\vec{x} = \int \psi(\vec{v})^* \phi(\vec{v}) d\vec{v} = \langle \psi, \phi \rangle.$$

As $SO(3)$ is a Lie group, we get an induced Lie algebra representation $\pi' : \mathfrak{so}(3) \rightarrow \mathfrak{u}(\mathcal{H}) : l \rightarrow \frac{d}{dt}\pi(e^{tl})|_{t=0}$. Recalling that $\mathfrak{u}()$ consists of skew adjoint matrices, we see that $L = i\hbar\pi'(l)$ is a hermitian operator on \mathcal{H} for any $x \in \mathfrak{so}(3)$ (The factor of \hbar is introduced so that the eigenvalues of L_j can be expressed in SI units).

A basis for $\mathfrak{so}(3)$ is given by

$$l_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad l_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad l_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

with corresponding observables $L_j = i\hbar\pi'(l_j)$. To begin with, we find an expression for L_j when using the standard basis for \mathbb{R}^3 . For a given function $f \in L^2(\mathbb{R}^3)$,

$$\begin{aligned} \pi'(l_1)f(\vec{x}) &= \frac{d}{dt}\pi(e^{tl_1})f(\vec{x})\Big|_{t=0} \\ &= \frac{d}{dt}f\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & \sin(t) \\ 0 & -\sin(t) & \cos(t) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right)\Big|_{t=0} \\ &= \frac{d}{dt}f(x_1, x_2 \cos(t) + x_3 \sin(t), -x_2 \sin(t) + x_3 \cos(t))\Big|_{t=0} \\ &= x_3 \frac{\partial f}{\partial x_2} - x_2 \frac{\partial f}{\partial x_3}. \end{aligned}$$

Similar calculations show that

$$\pi'(l_2)f(\vec{x}) = \left(x_3 \frac{\partial f}{\partial x_1} - x_1 \frac{\partial f}{\partial x_3}\right) \text{ and } \pi'(l_3)f(\vec{x}) = \left(x_2 \frac{\partial f}{\partial x_1} - x_1 \frac{\partial f}{\partial x_2}\right).$$

Therefore, expressing any $l = al_1 + bl_2 + cl_3 \in \mathfrak{so}(3)$ as a vector \vec{n} in three dimensions, we see that

$$i\hbar\pi'(l)f(\vec{x}) = \vec{n} \cdot (\vec{x} \times i\hbar\nabla f(\vec{x})).$$

Classically, the quantity $\vec{x} \times i\hbar\nabla f(\vec{x})$ represents the angular momentum of the wave f in three dimensions. Therefore, $L = i\hbar\pi'(l)$ must correspond to the operator which measures the component of angular momentum in the direction represented by \vec{n} .

Now that we have an understanding of what L measures, our next task is to find its eigenvectors and eigenvalues. Once this is done, the state of our

particle can be expanded in the resulting eigenbasis. This will allow us to assign a probability to each of the possible values of angular momentum.

In view of the Peter Weyl theorem, the compactness of $SO(3)$ implies that \mathcal{H} decomposes as a direct sum of finite dimensional vector spaces V_n . These are irreducible under $SO(3)$ and $\mathfrak{so}(3)$. As there is exactly one irreducible representation of $\mathfrak{so}(3)$ for each odd dimension, we know that

$$\mathcal{H} = V_1 \oplus V_3 \oplus V_5 \oplus \dots$$

Furthermore, we know that each V_{2l+1} is spanned by a basis of $2l+1$ functions $Y_{-l}^l, Y_{-l+1}^l, \dots, Y_{l-1}^l, Y_l^l$ where $L_3 Y_m^l = m Y_m^l$, $L_- Y_m^l = Y_{m-1}^l$ and $L_+ Y_m^l = Y_{m+1}^l$. Suppose that Y_l^l is the highest weight vector of V_{2l+1} . If we transform into spherical coordinates via

$$\begin{aligned} x_1 &= r \sin(\theta) \cos(\phi) \\ x_2 &= r \sin(\theta) \sin(\phi) \\ x_3 &= r \cos(\theta) \end{aligned}$$

then the action of $SO(3)$ will leave r invariant, so we can consider Y_m^l to be a function of just θ and ϕ . Now

$$\begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \phi & \cos \theta \cos \phi & -\sin \phi \\ \sin \theta \sin \phi & \cos \theta \sin \phi & \cos \phi \\ \cos \theta & -\sin \theta & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{1}{r} \frac{\partial}{\partial \theta} \\ \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \end{pmatrix},$$

so we can write

$$\begin{aligned} L_1 &= i\hbar \left(x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3} \right) = i\hbar \left(\sin(\phi) \frac{\partial}{\partial \theta} + \cot(\theta) \cos(\phi) \frac{\partial}{\partial \phi} \right) \\ L_2 &= i\hbar \left(x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3} \right) = i\hbar \left(-\cos(\phi) \frac{\partial}{\partial \theta} + \cot(\theta) \sin(\phi) \frac{\partial}{\partial \phi} \right) \\ L_3 &= i\hbar \left(x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2} \right) = -i\hbar \frac{\partial}{\partial \phi}. \end{aligned}$$

Further calculations show that

$$L_+ = \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot(\theta) \frac{\partial}{\partial \phi} \right) \text{ and } L_- = \hbar e^{-i\phi} \left(-\frac{\partial}{\partial \theta} + i \cot(\theta) \frac{\partial}{\partial \phi} \right).$$

Considering the highest weight vector $Y_l^l(\theta, \phi)$, we therefore have two differential equations:

$$L_3 Y_l^l(\theta, \phi) = -i\hbar \frac{\partial}{\partial \phi} Y_l^l(\theta, \phi) = l Y_l^l(\theta, \phi)$$

and

$$L_+ Y_l^l(\theta, \phi) = \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot(\theta) \frac{\partial}{\partial \phi} \right) Y_l^l(\theta, \phi) = 0.$$

Solving these equations, we see that

$$Y_l^l(\theta, \phi) = C_l e^{il\phi} \sin^l(\theta)$$

for some constant C_l . The lowering operator L_- can then be applied repeatedly to find each $Y_m^l = L_- Y_{m+1}^l$ for $m = l, l-1, \dots, l-1, -l$.

To summarize, we have now found an expression in spherical coordinates of each of the eigenfunctions of L_3 (Note that L_3 is special only because of the arbitrary way in which we transformed into spherical coordinates). We could therefore express any function in \mathcal{H} in terms of these eigenfunctions, allowing for a concrete description of the state of the particle.

In a spherically symmetric potential, the Hamiltonian operator H commutes with the action of L_3 and L_\pm . This implies that each of the eigenspaces of L_3 must be fixed by H . However, because every eigenspace of L_3 is one dimensional, they must also be energy eigenspaces. As H also commutes with L_\pm , we see that each V_l must in fact be an energy eigenspace. As such, we can label these spaces by their energy. This is the reason that there are exactly $2n+1$ possible values of the orbital angular momentum at energy level n .

5.5 Spin

It is also possible to model the state of a particle with a finite dimensional Hilbert space \mathbb{C}^n . In this situation, we imagine that a vector $\phi \in \mathcal{H}$ represents the ‘intrinsic properties’ of the particle. One such intrinsic property of all elementary particles is spin, which again arises from the action of $SO(3)$.

Suppose that \mathcal{H} is finite dimensional. According to an above section, we know that up to isomorphism, there is only one irreducible representation ρ of $\mathfrak{su}(2)$. This representation will give us self adjoint operators on \mathcal{H} in the following way: for any $X = aX_1 + bX_2 + cX_3 \in \mathfrak{su}(2)$

$$O(\hat{n}) = i\rho(\hat{n} \cdot X).$$

This operator

For example, if we wanted to describe the state of a single electron, we might choose \mathcal{H} to be the space of square integrable complex functions on \mathbb{R}^3 . In this situation, we imagine a state $\phi \in \mathcal{H}$ represents a probability cloud where $|\phi(\vec{x})|^2$ is the probability of being found around the point \vec{x} . Then $SO(3)$ acts on \mathcal{H} by

Now suppose that we have an observable O which is a function of direction \hat{n} . The key insight is that measuring $O(R\hat{n})$ is the same as measuring $O(\hat{n})$ in a reference frame that has been acted on by R . Therefore, for any $\phi, \psi \in \mathcal{H}$,

$$\langle \phi, O(R\hat{n})\psi \rangle = \langle \rho(R)\phi, O(\hat{n})\rho(R)\psi \rangle = \langle \phi, \rho(R)^\dagger O(\hat{n})\rho(R)\psi \rangle$$

Hence, any such observable must transform as

$$O(R\hat{n}) = \rho(R)^\dagger O(\hat{n})\rho(R) = \rho(R)^{-1} O(\hat{n})\rho(R)$$

as ρ is a unitary representation.

Then we will need to describe an operator $O(\hat{n})$ on the Hilbert space \mathcal{H} . From a rotated reference frame corresponds to $SO(3)$ acting on \hat{n} and \mathcal{H} simultaneously. As the physical properties of the system should be independent of the direction from which we view the system, if we measure the angular momentum along $R\hat{n}$

5.6 The periodic table

The total angular momentum is the sum of spin and orbital angular momentum. We see that this arises from . If a system begins in a given eigenspace, then any action of $\mathrm{SO}(3)$ will leave that eigenspace invariant, therefore ensuring that angular momentum is conserved. This is the essence of Noether's theorem.

6 Sources

<https://www.jmilne.org/math/CourseNotes/LAG.pdf>