MATH3001

Quantum Mechanics and Representation Theory

Harry Partridge

Term 1 2021

Contents

1 Introduction **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 Lie Theory **9**
3.1 Representations of Topological Groups 9
3.2 Lie groups 10
3.3 Lie Algebras 12
3.4 The Lie Algebra of a Lie group 13
3.5 Lie algebras of U(n), SL(n, $\mathbb{C}$) and SU(n) 15
3.6 Complex Irreps of $\mathfrak{sl}(2, \mathbb{C})$ and $\mathfrak{su}(2)$ 16
3.7 Complex Irreps of SO(3) and SU(2) 17

4 Quantum Mechanics **18**
4.1 The Quantum Formalism 18
4.2 Measurement 20
4.3 Lie Theory, Representations in Quantum Mechanics 21
4.4 Orbital Angular Momentum 21
4.5 Spin 24
4.6 Tensor products and the periodic table 25

References **25**

1 Introduction

The purpose of this report is to explore representation theory and 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 $e^{(A^\dagger)} = \left(e^A\right)^\dagger = e^{-A}$, 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 $(\pi(R)f)(x) = f(R^{-1}x)$. This generates the orbital angular momentum operators.
2. Say $\mathcal{H}$ is $\mathbb{C}^n$. Then $SU(2)$ acts on $\mathcal{H}$ by $(\pi(R)f)(x) = f(R^{-1}x)$. This generates the spin angular momentum operators.
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

2.1 Basic Definitions

There are several equivalent ways to define a representation of a group [1]. The first is via a homomorphism.

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

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

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 a 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 $\theta : V \to \mathbb{F}^n$ 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 $\rho : G \to GL(V)$ 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 $\lambda^n = 1$. □

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 $\rho : G \to GL(V)$ 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. □

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 $\rho : G \to GL(V)$ 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 $\rho_{|W} : G \to GL(V)$ 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 $\rho : G \to GL(V)$ 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 $\rho = \rho_W \oplus \rho_W$.

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 $f : V \to W$ 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 **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 $1_G(g) = (id : \mathbb{F} \to \mathbb{F})$ 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 $\rho(g) = \mathrm{id}_V$.

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

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

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

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

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

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

Therefore $\mathrm{im}(\bar{f}) = W$, 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 $h \in G$,

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

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

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

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. □

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 $f : V \to W$ 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 $\mathrm{im}(f) = W$, 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 $f = \lambda I$. □

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 $\rho : G \to GL(V)$ 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 $\rho(h) \in \mathrm{End}(V)$ 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. □

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

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

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

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 $\mathrm{tr}(R_1) = \mathrm{tr}(R_2)$. 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 $\rho : G \to GL(V)$, $\mathrm{tr}(\rho(g))$ is well defined for every $g \in G$. For every representation we therefore have an associated function $\chi : G \to GL(V) \to \mathbb{F}$ given by $\chi = \mathrm{tr} \circ \rho$. This is known as the character of G.

Definition (Character)**.** Suppose $\rho : G \to GL(V)$ 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** $\chi : G \to \mathbb{F}$ defined by $\chi(g) = \mathrm{tr}(\rho(g))$ 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 $\chi : G \to GL(V) \to \mathbb{F}$ 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 $\chi(g) = \mathrm{tr}(\mathrm{id}_V) = \dim(V) = n$.

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 $\rho(g) = \lambda\,\mathrm{id}_V$ 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 $\chi(g) = \operatorname{tr}(f \circ \rho'(g) \circ f^{-1}) = \chi'(g)$. □

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:

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

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, $\chi(ghg^{-1}) = \operatorname{tr}\left(\rho(g)\rho(h)\rho(g)^{-1}\right) = \chi(h)$. □

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 $(\operatorname{char}(\mathbb{F}), |G|) = 1$*?*

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. □

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. □

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

Proof. TO DO. □

Corollary (Irreduciblity 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_0. Then exactly one $m_i = 1$, and the rest are zero, so χ is irreducible. □

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 Lie Theory

3.1 Representations of Topological Groups

A topological group is a group that has been endowed with a topology. Most of the definitions regarding representations of topological groups are very similar to those for finite groups, but typically also require some kind of continuity. The following definitions are given by [2].

Definition (Topological group)**.**

1. A group G endowed with a Hausdorff topology is a **topological group** if
   (a) The left and right multiplication maps $L_g : G \to G : h \mapsto gh$ and $R_g : G \to G : h \mapsto hg$ are continuous for every g.
   (b) The inverse map $\text{inv} : G \to G : g \mapsto g^{-1}$ is continuous.
2. A **morphism** between topological groups is a continuous group homomorphism.

Primarily, we are interested in the case where G is compact under the topology τ. Some examples of compact topological groups include all finite groups with discrete topology, discrete groups such as $\mathbb{Z}$ (also with discrete topology), and Lie groups.

Definition (G-space)**.**

1. A locally compact topological space X is a **G-space** if there is a continuous map $X \to X$ given by $x \mapsto gx$ for every $g \in G$ and
   (a) $g_1(g_2x) = (g_1g_2)x$ for all $g_1, g_2 \in G$
   (b) $ex = x$ where e is the identity in G.
2. A **morphism** between two G-spaces X and Y is a continuous map $\phi : X \to Y$ satisfying $\phi(gx) = g\phi(x)$ for all $g \in G, x \in X$.

Representations of G arise when X is also endowed with a vector space structure.

Definition (Topological vector space)**.** A **topological vector space** is a vector space equipped with a Hausdorff topology for which addition and scalar multiplication are continuous.

Example. Suppose that V is a finite dimensional vector space. Then the standard topology inherited via a linear isomorphism $V \simeq R^n$ is the unique topology which makes V into a topological vector space. [2]

For our purposes, the most important example of a topological vector space is a Hilbert space.

Definition (Hilbert Space)**.** Suppose V is a vector space endowed with an inner product $\langle \cdot, \cdot \rangle$. Then V can be made into a topological vector space by the the topology induced by the metric $d(x, y) = \sqrt{\langle x, y \rangle}$. V is then a **Hilbert space** if it is complete with respect to the metric d.

For any topological vector space V, we can define $GL(V)$ to be the set of continuous automorphisms on V. This set is itself a finite dimensional vector space, and is therefore made into a topological vector space in the canonical way.

As is the case with finite groups, representations of topological groups can be viewed as either homomorphisms, linear actions or modules. We give one definition here; the other viewpoints relate in the same way as before.

Definition (G-representation)**.**

1. Suppose G is a topological group. A **representation** of G on a topological vector space V is a continuous homomorphism $\pi : G \to GL(V)$. The induced linear action of G on V is given by $gv = \pi(g)v$.

2. A **morphism** between G-representations π_1 and π_2 on V_1 and V_2 respectively is continuous linear map $T : V_1 \to V_2$ such that $T \circ \pi_1(g) = \pi_1(g) \circ T$ for all $g \in G$. It is clear that this is a G-space morphism when V_1 and V_2 are considered G-spaces under the actions induced by π_1 and π_2.

If V is a Hilbert space, we have a particularly important kind of representation.

Definition (Unitary representation)**.** A representation π of G on a Hilbert space $\mathcal{H}$ is **unitary** if $\langle gv, gw \rangle = \langle v, w \rangle$ for all $v, w \in V$ and $g \in G$.

3.2 Lie groups

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 $\phi_\alpha : V_\alpha \to U_\alpha$, 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

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

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 $f : M \to N$ 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.

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

Definition (Lie group)**.**

1. A group G endowed with a Hausdorff topology τ and smooth structure (V_α, ϕ_α) is a **Lie group** iff
   (a) The left and right multiplication maps $L_g : G \to G : h \mapsto gh$ and $R_g : G \to G : h \mapsto hg$ are smooth for every g.
   (b) The inverse map $\text{inv} : G \to G : g \mapsto g^{-1}$ is smooth.
2. A **morphism** between Lie groups is a smooth group homomorphism.

Noting that L_g and R_g have smooth inverses $L_{g^{-1}}$ and $R_{g^{-1}}$, we see that these maps are in fact diffeomorphisms.

The first example of a Lie group is the general linear group $GL(V)$, where the manifold structure inherited from $\mathbb{R}^{n^2}$. This allows us to define representations of a Lie group.

Definition (Lie Group Representation)**.** For a given Lie group G, a representation on a vector space V is a Lie group morphism $\pi : G \to GL(V)$.

There are several other important examples of Lie groups. In each of the following, the manifold structure is inherited from $\mathbb{C}^{n^2}$ or $\mathbb{R}^{n^2}$.

1. $U(n)$ is the group of $n \times n$ complex unitary matrices
$$U(n) = \{X \in GL(n, \mathbb{C}) : XX^\dagger = 1\}.$$
2. $O(n)$ is the group of $n \times n$ real orthogonal matrices
$$U(n) = \{X \in GL(n, \mathbb{R}) : XX^T = 1\}.$$
3. $SL(n, \mathbb{C})$ is the group of $n \times n$ complex matrices with determinant 1:
$$SL(n, \mathbb{C}) = \{X \in GL(n, \mathbb{C}) : \det(X) = 1\}.$$
4. $SL(n, \mathbb{R})$ is the group of $n \times n$ real matrices with determinant 1:
$$SL(n, \mathbb{C}) = \{X \in GL(n, \mathbb{R}) : \det(X) = 1\}.$$
5. $SU(n)$ is the intersection of $SL(n, \mathbb{C})$ and $U(n)$:
$$SU(n) = \{X \in GL(n, \mathbb{C}) : XX^\dagger = 1, \det(X) = 1\}.$$

6. $SO(n)$ is the intersection of $SL(n, \mathbb{R})$ and $O(n)$:

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

Although Lie groups can often be very large, the following theorem shows that they can be characterised by an arbitrarily small open set.

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 $G = S$. □

It turns out that we can extend this idea further. The region around the identity of the Lie group can in fact be 'Linearised' to form a Lie algebra that completely characterizes the connected component of the identity. This is the focus of the following two sections.

3.3 Lie Algebras

We begin with the definition of a Lie algebra.

Definition (Lie algebra)**.**

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

   (a) The bracket operation is bilinear.

   (b) $[x, x] = 0$ for all $x \in V$.

   (c) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$

   The third requirement is known as the Jacobi identity.

2. A **morphism** between Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is a linear function $\phi : \mathfrak{g} \to \mathfrak{h}$ such that

$$\phi([X, Y]) = [\phi(X), \phi(Y)].$$

The first thing to note is that applying requirements (a) and (b) to the bracket $[x + y, x + y]$, 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 $x, y \in V$.

Representations are then defined in the usual way.

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 morphism $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$.

3.4 The Lie Algebra of a Lie group

This section illustrates how a Lie algebra arises from the tangent space at the identity of a Lie group.

For any $g \in G$, we define the conjugation action $\Psi_g : G \to G$ by $\Psi_g(h) = L_g \circ R_g(h) = ghg^{-1}$. Since Ψ_g fixes the identity e in G, the derivative $(d\Psi_g)_e : T_eG \to T_eG$ is an endomorphism on T_eG. As the composition of diffeomorphisms, Ψ_g is a diffeomorphism. Therefore, $d\Psi_g^{-1}$ is non singular, so $(d\Psi_g)_e^{-1}$ exists. Therefore $(d\Psi_g)_e$ is actually an automorphism on T_eG. This motivates the following definition.

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

Now $\mathrm{Aut}(T_eG)$ is an open subset of $\mathrm{End}(T_eG)$, so the tangent space of $\mathrm{Aut}(T_eG)$ at the identity $I = \mathrm{id}_{T_eG}$ can be naturally identified with $\mathrm{End}(T_eG)$. As such, the derivative of Ad is a map

$$\mathrm{ad} = (d\,\mathrm{Ad})_I : T_eG \to \mathrm{End}(T_eG).$$

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

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

Instead of providing a complete proof, we show that this definition corresponds to the commutator when $G = GL(\mathbb{R}^n)$.

Proof. The first thing to note is that when $G = GL(\mathbb{R}^n)$, the exponential map $\exp : T_eG \to G : X \mapsto \sum_{n=0}^{\infty} X^n/n!$ is well defined. Therefore, for any $g \in G$,

and $X \in T_eG$,

$$\begin{aligned}
\mathrm{Ad}(g)(X) &= d\Psi_g|_e(X) \\
&= \frac{d}{dt}\Psi_g(\exp(tX))\Big|_{t=0} \\
&= \frac{d}{dt} g\exp(tX)g^{-1}\Big|_{t=0} \\
&= gA\exp(tX)g^{-1}|_{t=0} \\
&= gXg^{-1}.
\end{aligned}$$

We then see that for any $X, Y \in T_eG$,

$$\begin{aligned}
\mathrm{ad}(X)(Y) &= d\,\mathrm{Ad}|_e(X)(Y) \\
&= \frac{d}{dt}\,\mathrm{Ad}(\exp(tX))(Y)\Big|_{t=0} \\
&= \frac{d}{dt}\exp(tX)Y\exp(tX)^{-1}\Big|_{t=0} \\
&= \frac{d}{dt}\exp(tX)Y\exp(-tX)\Big|_{t=0} \\
&= X\exp(tX)Y\exp(-tX) - \exp(tX)YX\exp(-tX)|_{t=0} \\
&= XY - YX.
\end{aligned}$$

□

It turns out that we can define an exponential map even in more general situations. Suppose G is a Lie group with Lie algebra $\mathfrak{g}$. For any $X \in \mathfrak{g}$, we can define a vector field v_X by $V_X(g) = dL_g|_e(X)$, where L_g is the left action of g. Integrating v_X, we get a curve ϕ from an open neighborhood of 0 I to G. According to [3], ϕ is a homomorphism of Lie groups, so it extends to the entire real line. We can now define the exponential map.

Definition (Exponential map)**.** Suppose G is a Lie group with Lie algebra $\mathfrak{g}$. The exponential map $\exp : \mathfrak{g} \to G$ is defined by

$$\exp(X) = \phi_X(1),$$

where ϕ_X is the integral curve described above.

According to [3], there is also a one to one correspondence between morphisms of two Lie groups and morphisms of their associated Lie algebras.

Proposition. Suppose G and H are Lie groups with G simply connected. A morphism $\pi : G \to H$ is uniquely determined by its differential $\pi' := d\pi|_e : \mathfrak{g} \to \mathfrak{h}$. Furthermore, these maps commute with the exponential:

$$\pi(\exp(X)) = \exp(\pi'(X))$$

for any $X \in \mathfrak{g}$.

This is particularly useful because it implies that a Lie group representation π induces an associated Lie algebra representation π'. We will use this many times in the section on Quantum mechanics.

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

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 $X + X^\dagger = 0$. Note that this argument also holds for infinite dimensional unitary groups. In any case, we see that 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. □

In the case of $SL(n, \mathbb{C})$, we can again 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 $\det(e^X) = e^{\mathrm{tr}(X)}$, we have that $\mathrm{tr}(X) = 0$.

Proposition. The Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ 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 $\mathfrak{sl}(n, \mathbb{C})$ because $\mathfrak{sl}(n, \mathbb{C})$ is a subset with the same dimension. □

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

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

In the case where $n = 2$, there is a nice relationship between $\mathfrak{su}(n)$ and $\mathfrak{sl}(n, \mathbb{C})$. First of all, we can write

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

On the other hand, since $\mathfrak{sl}(2, \mathbb{C})$ is just the set of traceless 2×2 complex matrices, so an element of $\mathfrak{sl}(2, \mathbb{C})$ is of the form

$$\begin{pmatrix} x - iy & a + ib \\ c + id & -x + iy \end{pmatrix} = \begin{pmatrix} -iy & s + iu \\ -s + iu & iy \end{pmatrix} + i \begin{pmatrix} -ix & v + it \\ -v + it & ix \end{pmatrix}$$

where $s = \frac{1}{2}(a-c)$, $t = -\frac{1}{2}(a+c)$, $u = \frac{1}{2}(b+d)$ and $v = \frac{1}{2}(b-d)$. Hence, it is not hard to see that when viewed as a vector space over $\mathbb{R}$, $\mathfrak{sl}(2,\mathbb{C}) \simeq \mathfrak{su}(2) \oplus i\mathfrak{su}(2)$.

To make this situation a little clearer, we give a basis for $\mathfrak{su}(2)$ that relates nicely to a basis for $\mathfrak{sl}(2,\mathbb{C})$. It is not too hard to see that $\mathfrak{su}(2) = \operatorname{span}_{\mathbb{R}}\{X_1, X_2, X_3\}$, where

$$X_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad X_3 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

These matrices obey the commutation relations

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

We can then define $e = \frac{1}{2}(X_2 - iX_1)$, $f = -\frac{1}{2}(X_2 + iX_1)$ and $h = iX_3$ such that

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

These matrices obey the commutation relations

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

Furthermore, $\mathfrak{sl}(2,\mathbb{C}) = \operatorname{span}_{\mathbb{R}}\{X_1, X_2, X_3, iX_1, iX_2, iX_3\} = \operatorname{span}_{\mathbb{C}}\{e, f, h\}$.

3.6 Complex Irreps of $\mathfrak{sl}(2,\mathbb{C})$ and $\mathfrak{su}(2)$

Suppose that $\mathfrak{sl}(2,\mathbb{C})$ acts linearly on a complex 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 $v \in V_\lambda$, $fv \in V_{\lambda-2}$ and $ev \in V_{\lambda+2}$.

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

As $\mathbb{C}$ is algebraically closed, h must have at least one eigenvalue with an eigenvector u. Therefore, by finite dimensionality, there must be an eigenvector v of h such that $ev = 0$. We say that such an eigenvector is a **maximal vector** of V and that the corresponding eigenvalue λ is a **maximal weight**. Similarly, we say that an eigenvector w of h such that $fw = 0$ is a **minimal vector**, with **minimal weight**.

Theorem (Irreducible representations of $\mathfrak{sl}(2,\mathbb{C})$)**.** Let V be an n dimensional complex vector space with an irreducible representation π of $\mathfrak{sl}(2,\mathbb{C})$. Then V decomposes into a sum of one dimensional weight spaces

$$V_{-l}, V_{-l+2}, \ldots, V_{l-2}, V_l,$$

where $l = n - 1$ is an integer.

Proof. Repeatedly applying $\pi(f)$ to the highest weight vector v_λ, we get a set β of $k+1$ eigenvectors of $\pi(h)$:

$$\beta = \{v_\lambda, v_{\lambda-2}, v_{\lambda-4}, \dots, v_{\lambda-2k} = w\}$$

where $v_{\lambda-2j} = \pi(f)^j v_\lambda$ and w is the minimal vector. Let $W = \mathrm{span}_{\mathbb{C}}(\beta) \subseteq V$. Then we claim that W is invariant.

As $\pi(f)(\beta) = \{v_{\lambda-2}, v_{\lambda-4}, \dots, w\} \subset \beta$, W is invariant under the action of $\pi(f)$. Similarly, β is a set of eigenvectors of $\pi(h)$, so W must be $\pi(h)$ invariant. We now use induction to show that for all $j \leq k+1$, (where $v_{\lambda-2(k+1)} = 0$)

$$\pi(e)v_{\lambda-2j} = j(\lambda - j + 1)v_{\lambda-2(j-1)}.$$

Since v_λ is the highest weight vector, $\pi(e)v_\lambda = 0$, so the statement holds for $j = 0$. Assuming validity for $j \leq k$, we see that

$$\begin{aligned}
\pi(e)v_{\lambda-2(j+1)} &= \pi(e)\pi(f)v_{\lambda-2j} \\
&= (\pi([e,f]) + \pi(f)\pi(e))v_{\lambda-2j} \\
&= (\pi(2h) + \pi(f)\pi(e))v_{\lambda-2j} \\
&= ((\lambda-2j)v_{\lambda-2j} + \pi(f)j(\lambda-j+1)v_{\lambda-2(j-1)} \\
&= ((\lambda-2j) + j(\lambda-j+1))v_{\lambda-2j} \\
&= (j+1)(\lambda-(j+1)+1)v_{\lambda-2((j+1)-1)}.
\end{aligned}$$

Hence W must also be $\pi(e)$ invariant, and therefore also invariant under the action of any $x \in \mathfrak{sl}(2,\mathbb{C})$. However, because V is irreducible, this implies that $W = V$ and $k = l$.

Finally, note that $\pi(e)v_{\lambda-2(l+1)} = \pi(e)\pi(f)w = 0$. Therefore,

$$(l+1)(\lambda-(l+1)+1) = (l+1)(\lambda-l) = 0,$$

so $\lambda = l$, and the theorem is proved. □

This classification carries over to $\mathfrak{su}(2)$ in the following way. If π is a finite dimensional complex irreducible representation of $\mathfrak{su}(2)$ on V, then π^* defined by $\pi^*(h) = i\pi(X_3), \pi^*(e) = \frac{1}{2}\pi(X_2) - \frac{i}{2}\pi(X_1)$, and $\pi^*(f) = -\frac{1}{2}\pi(X_2) - \frac{i}{2}\pi(X_1)$ is an irrep of $\mathfrak{sl}(2,\mathbb{C})$. Therefore, V decomposes in the way described above, and each of the weight spaces V_m are also eigenspaces of $\pi(X_3)$.

3.7 Complex Irreps of SO(3) and SU(2)

Now that we have classified the complex irreps of $\mathfrak{su}(2)$, we can use the exponential map to find the complex irreps of $SU(2)$. Explicitly, since $SU(2)$ is simply connected, any representation π on V is completely determined by its differential π', which is a representation of $\mathfrak{su}(2)$. Therefore, V must decompose in the way described in the above section.

We now want to find the irreps of $SO(3)$. A basis for $\mathfrak{so}(3)$ is given by

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

These matrices obey the commutation relations

$$[l_1, l_2] = l_3, \qquad [l_2, l_3] = l_1, \qquad [l_3, l_1] = l_2.$$

Noting that this is the same set of relations as for $\mathfrak{su}(2)$, we can define a Lie algebra isomorphism $\pi' : \mathfrak{so}(3) \to \mathfrak{su}(2) : X_j \mapsto l_j$. This means that the irreps of $\mathfrak{so}(3)$ are the same as those for $\mathfrak{su}(2)$. Importantly however, this does not mean that the irreps of $SO(3)$ are the same as those for $SU(2)$. Using the exponential map, we can define a Lie group homomorphism $\pi : SO(3) \to SU(2)$ by

$$\pi \exp\left(\vec{n} \cdot \vec{l}\right) = \exp\left(\vec{n} \cdot \pi'\left(\vec{l}\right)\right),$$

It turns out that there is a nice relationship between $SO(3)$, $SU(2)$ and the unit quaternions. Any element $g \in SU(2)$ has the form

$$g = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$$

where $z, w \in \mathbb{C}$ and $|z|^2 + |w|^2 = 1$. Writing $z = a + ib, w = c + id$, we see that $a^2 + b^2 + c^2 + d^2 = 1$. This equation defines the three sphere S^3 in $\mathbb{R}^4$, which is topologically equivalent to the unit sphere in the quaternions [4]. In fact, we can define $\phi : SU(2) \to \mathbb{H}$, by

$$\phi\begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix} = a + bi + cj + dk.$$

Routine calculations show that this is an isomorphism.

For any point $(x, y, z) \in \mathbb{R}^3$, we can associate an imaginary quaternion $q = xi + yj + zk$. Suppose that $g \in SU(2)$ has $\phi(g)$ as its associated quaternion. Then more calculations show that $\phi(g)q\phi(g)^{-1} = ui + vj + wk$ is another imaginary quaternion, which can be associated with a point (u, v, w) in $\mathbb{R}^3$. According to [4], this gives us a $2:1$ surjective homomorphism $\Phi : SU(2) \to SO(3)$. In particular, note that $\Phi(-g)(q) = \phi(-g)q\phi(-g)^{-1} = \phi(g)q\phi(g)^{-1} = \Phi(g)(q)$, so $\Phi(-g) = \Phi(g)$. In this situation we say that $SU(2)$ is a double cover of $SO(3)$.

The presence of Φ allows us to deduce the irreducible representations of $SO(3)$.

The net result of this is that there is exactly one complex irreducible representation of $SO(3)$ for each odd dimension.

4 Quantum Mechanics

4.1 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 the energy of a particle would be given as $E : S \to \mathbb{R} : E(x, v) = mv^2/2$.

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.

4.2 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

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

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 eignstates 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.

4.3 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 a outline of how this works, and the remaining sections examine several examples.

Suppose that a compact group G has a strongly continuous unitary representation $\pi : G \to U(\mathcal{H})$ 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.

4.4 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 $\pi : SO(3) \to GL(H)$ 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 $R \in SO(3)$, a change of variables $\vec{x} \mapsto R\vec{x}$ has a Jacobian determinant of 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 $\pi' : \mathfrak{so}(3) \to \mathfrak{u}(\mathcal{H}) : l \to \frac{d}{dt}\pi(e^{tl})\big|_{t=0}$. Recalling that $\mathfrak{u}(\mathcal{H})$ 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).

Recall that a basis for $\mathfrak{so}(3)$ is given by matrices l_1, l_2, and l_3, 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 $f \in L^2(\mathbb{R}^3)$,

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

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 corollary to 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 which are irreducible under $SO(3)$. As there is exactly one irreducible representation of $SO(3)$ 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 $Y^l_{-l}, Y^l_{-l+1}, \dots, Y^l_{l-1}, Y^l_l$ where $L_3 Y^l_m = mY^l_m$, $L_- Y^l_m = Y^l_{m-1}$ and $L_+ Y^l_m = Y^l_{m+1}$. 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^l_m to be a function of just θ and ϕ. Now

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

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:

$$L_3 Y^l_l(\theta,\phi) = --i\hbar\frac{\partial}{\partial\phi}Y^l_l(\theta,\phi) = lY^l_l(\theta,\phi)$$

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

$$Y^l_l(\theta,\phi) = C_{ll}e^{il\phi}\sin^l(\theta)$$

for some constant C_{ll}. The lowering operator L_- can then be applied repeatedly to find each $Y^l_m = L_- Y^l_{m+1}$ 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.

4.5 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 arises from the action of $SU(2)$.

Suppose that $\mathcal{H} = \mathbb{C}^n$. We know than any representation of $SU(2)$ decomposes as a sum of irreducible representations, so we restrict ourselves to the case where $\pi : SU(2) \to GL(\mathcal{H})$ is irreducible. We also want π to represent a symmetry of $\mathcal{H}$, so we choose π to be unitary. As usual, we also get a Lie algebra representation $\pi' : \mathfrak{su}(2) \to \mathfrak{gl}(\mathcal{H})$. Fortunately, we have already classified such representations: we know that $\mathcal{H}$ must have a basis of n eigenvectors v_m such that $i\pi'(X_3)v_m = mv_m$ for $m = -n/2, -n/2+1, \ldots, n/2-1, n/2$.

Once agin, the operators $\pi'(X_j)$ are skew adjoint, so $i\hbar\pi'(X_j)$ is an observable. It turns out that these observables measure 'spin', which is a quantum property for which there is no classical analogue (although it bears some similarities to the angular momentum of a three dimensional object spinning about an internal axis).

In general, for any $X = aX_1 + bX_2 + cX_3 \in \mathfrak{su}(2)$, we can write $\vec{n} = (a, b, c)$, $S = i\hbar(\pi'())$, and then

$$O(\hat{n}) = \hat{n} \cdot \pi'(X).$$

This operator

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 $\phi, \psi \in \mathcal{H}$,

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

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 $R\hat{n}$

4.6 Tensor products and the periodic table

In the proceeding two sections, we examined an infinite dimensional Hilbert space $L^2(\mathbb{R}^3)$, which is used to model a particle in space, and a finite dimensional Hilbert space $\mathbb{C}^n$ which is used to model the intrinsic properties of a particle. We would now like to be able to do both at once. This is done by constructing a new Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^n$, which is used to model particles in three dimensional space with multiple possible spin states.

The most important example of this is that of the electron. We know that the electron has two possible spin states, so we use $\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. In this situation, the operators which we had constructed on $L^2(\mathbb{R}^3)$ will leave the $\mathbb{C}^2$ subspace invariant and vise versa. The result of this is that the Hamiltonian and angular momentum operators commute with the spin operators, allowing us to break up the entire state into one dimensional eigenspaces labeled by energy, angular momentum and spin.

These states are the source of many basic facts about the periodic table. For example, the number of electrons per orbital is $2, 6, 10, 14, \ldots$. These numbers correspond to the fact that there are $1, 3, 5, 7, \ldots$ possible values of the angular momentum, and two possible spin states for each energy level.

References

[1] D. Chua, *Representation Theory - 2016 lectures by S. Martin.* 2016.

[2] A. Din, "Representation theory of compact groups." `http://www.math.caltech.edu/~2016-17/2term/ma145b/notes/RepCptGroups.pdf`, 2017.

[3] W. Fulton and J. Harris, *Representation theory: A first course.* Springer-Verlag, 1991.

[4] Wikipedia contributors, "Special unitary group," 2020. [Online; accessed April-2020].