An Introduction to Lie Theory and Applications

by

Anthony Dickson

Submitted in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF SCIENCE

in the

Mathematics

Program

YOUNGSTOWN STATE UNIVERSITY

May, 2021

An Introduction to Lie Theory and Applications

Anthony Dickson

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Anthony Dickson, Student Date

Approvals:

Dr. Thomas Madsen, Thesis Advisor Date

Dr. Tom Wakefield, Committee Member Date

Dr. Michael Crescimanno, Committee Member Date

Dr. Salvatore A. Sanders, Dean of Graduate Studies Date

ABSTRACT

This thesis goes over various topics of Lie theory and is meant as an introduction for those who have never studied the subject. This is accomplished by first reviewing necessary background material, including topics from linear algebra and topology (where no background knowledge is assumed), before proceeding to the main subject. We start by discussing the matrix Lie group before discussing what a Lie algebra is. We define a Lie algebra as a vector space with additional requirements before defining it again with relation to a Lie group. We then explore various properties and examples of this. Next, we turn to representation theory and how it can be applied to Lie theory, including the important subject of roots and weights. Finally we finish by briefly going over some connections Lie theory has with physics (no background physics knowledge is assumed of the reader). Overall, the main goal of this thesis is to be an accessible starting point for someone who has a strong background in linear and abstract algebra, but has never studied Lie theory.

ACKNOWLEDGEMENTS

First I would like to thank my committee members, starting with Dr. Thomas Madsen. I first had him as my Abstract Algebra 1 instructor and his enthusiasm inspired me to further pursue algebra (as well as have him for an additional two classes, an advisor for my undergraduate thesis, and my primary advisor for this thesis). Unsurprisingly, this past year has been very challenging, but Dr. Madsen never pressured me into working more on the thesis than I already was. If I needed to cancel a meeting for a week, I could do that. If I needed more time to look at something before showing it to him, he was always okay with it. He never made me feel worse or like I was being lazy, and it's really awesome to have an advisor who trusts you to go at your own pace. Dr. Madsen's youthfully exuberant personality makes every conversation a joy. He radiates fun and that makes every meeting I have with him a blessing. Having a mentor whose research interest I enjoy is wonderful, but having a mentor whose personality makes me smile and laugh is the greatest. My goal as a future educator is to be as fun as him for anyone I work with – hopefully I can achieve that goal one day. Thank you.

I would like to thank Dr. Tom Wakefield. I have had him as an instructor for six different mathematics courses and he has been my academic advisor for the past five years. He has helped me excel in academia in ways I never dreamed of before, including encouraging me to join in our local Pi Mu Epsilon chapter, attend multiple conferences, and participate in an REU experience. I will forever be grateful for his guidance. In addition to being an extraordinary advisor, Dr. Wakefield is an even more extraordinary person. He responds to emails faster than a speeding bullet, always has a smile or laugh to share with the room, and it is impossible to leave a meeting with him and not have your day significantly improved. His kind nature makes him easygoing and accessible for students. In short, Dr. Wakefield is the coolest. Thank you.

I would like to thank Dr. Michael Crescimanno, my final committee member, whose physics expertise was helpful in making sure I (a mathematician with little physics background) did not say anything incorrect. His excitement in joining my committee and willingness to answer any questions I had was greatly appreciated. Thank you.

I would like to thank Dr. Kerns, the graduate coordinator, who helped answer any questions I had about the program, made sure I was on top of things, and helped with the formatting of this thesis. Thank you.

I would like to thank the entire Department of Mathematics and Statistics at Youngstown State University for providing a second-home for me and a place where I could not just intellectually thrive with a strong and caring faculty, but also socially thrive with some of the grooviest students. Thank you.

Finally, I would like to thank some of my closest friends for their support: Jonathan - for always making me smile with his silliness and infectious sense of humor, Kevin - for always going with the flow and saying hello, and Payton - whose contribution was invaluable and who I will mention again in the Introduction. To all these wonderful chaps: thank you.

I am sure I am forgetting people to thank, and so to everyone else: thank you. This thesis would not have been possible without my support system.

Contents

1 Introduction **1**

2 Background Material **3**

2.1 Linear Algebra 3

2.2 Dual Space 8

2.3 Tensor Products 9

2.4 Topology 12

3 Introduction to Lie Theory **15**

3.1 Lie Groups and the Exponential 15

3.2 Lie Algebras 20

3.3 Connecting Lie Groups and Lie Algebras 26

3.4 Examples of Lie Algebras 28

3.5 Lie Group and Lie Algebra Properties 33

4 Representation Theory **41**

4.1 Basics 41

4.2 Representations of Lie Groups and Algebras 42

4.3 Representations of Homogeneous Polynomials 48

4.4 Tensor Products of Representations 52

4.5 Dual Representations 54

4.6 Representations of $Lie(SL_2(\mathbb{C}))$ 55

4.7 Roots and Weights 58

4.8 Representations of $Lie(SL_3(\mathbb{C}))$ 60

4.9 The Weyl Group 66

5 Physics Applications **72**

5.1 The Rotation Group 72

5.2 The Physics Approach 73

5.3 Example with Highest Weight of $\frac{1}{2}$ 78

5.4 Combining Particles 81

6 Conclusion **83**

Bibliography **83**

Chapter 1

Introduction

The primary goal of this thesis is to provide a clear introduction to Lie theory for those who have not studied it, but have a suitable background (a solid understanding of linear and abstract algebra). No knowledge of representation theory is assumed, but this is an important topic as well. We also use some results from topology, but again, no prerequisite knowledge is presumed. We will review important background necessary to understand new topics. As part of this goal, we also want to use this knowledge to observe some interesting connections between Lie theory and physics. No physics knowledge is expected of the reader.

This thesis started as a project during the summer of 2020 between Payton Linton and myself. We worked together to understand Lie theory and he was an invaluable source to bounce ideas off of. Being a physics major, he was also very helpful with increasing my understanding of the physics contained here. Finally, he helped type the sections on roots and weights, and he was always available to help proofread the thesis. His help is greatly appreciated and I could not have done this without him.

For sources, we primarily used Brian C. Hall's *Lie Groups, Lie Algebras, and representations* (2015) [4]. Many of our definitions and theorems come from this source, and it was one of my goals to go through parts of this book and make it as accessible as possible for those with a less sophisticated background. In particular, we studied parts of the first four chapters, as well as Chapter 6, when studying Lie theory. For the Physics Applications chapter, we looked closely at Chapter 4: Group Theory and the Quark Model from Ta-Pei Cheng and Ling-Fong Li's *Gauge theory of elementary particle physics* [1]. Finally, we also consulted R. Shankar's *Principles of quantum mechanics* (2008) [5] and Woit's *Quantum theory, groups and representations* (2017) [7] to help digest some of the physics material.

The thesis is structured into five chapters. The first is this introduction, while the second contains some important background material (such as topics from linear algebra and elementary topology) deemed necessary to understand Lie theory. The third chapter contains the essence of Lie theory (including Lie groups and Lie algebras), while the fourth chapters uses that in a representation theory setting. Chapter 5 applies the concepts we have discussed to provide interesting physics connections. We then finish the thesis with some concluding remarks.

Without further ado, we can now start Chapter 2 and dive into some linear algebra.

Chapter 2

Background Material

This chapter will cover results from subjects such as linear algebra and topology, as well as detail some of the notation we will be using throughout this thesis.

2.1 Linear Algebra

Before we begin, here are some important notes:

- For the entirety of this thesis, all vector spaces we are considering will be finite-dimensional.
- We denote the **trace** of a matrix A by $tr(A)$.
- We denote the **transpose** of a matrix A by A^T.
- We denote the **adjoint**, or complex conjugate transpose, of a matrix A by A^*.
- We denote the $n \times n$ **identity matrix** by I_n.

Now we can begin with some important definitions from linear algebra.

Definition 2.1.1. *The* ***general linear group*** *of degree n over a field F is defined by $GL_n(F) = \{A \in M_n(F) : det(A) \neq 0\}$. Since $det(A) \neq 0 \iff A$ is invertible, this is also known as the set of $n \times n$ invertible matrices.*

Note that $GL_n(F)$ is a group (hence its name) and that examples of fields F that we can use are the complex field, $\mathbb{C}$, and the field of real numbers, $\mathbb{R}$. Also note that for a finite-dimensional vector space V, $GL(V) \cong GL_n(F)$.

Definition 2.1.2. *The **special linear group** of degree n over a field F is defined by $SL_n(F) = \{A \in M_n(F) : det(A) = 1\}$. Since $det(A) = 1 \implies det(A) \neq 0$, the special linear group is a subgroup of the general linear group.*

Definition 2.1.3. *The **orthogonal group** of degree n over a field F is defined by $O(n) = \{A \in GL_n(F) : A^{-1} = A^T\}$. Similarly, we can define the **special orthogonal group** of degree n over a field F as $SO(n) = \{A \in SL_n(F) : A^{-1} = A^T\}$.*

Definition 2.1.4. *The **unitary group** of degree n over a field F is defined by $U(n) = \{A \in GL_n(F) : A^{-1} = A^*\}$. Note that if A is a real matrix, then the complex conjugate transpose would just be the transpose, and so we would have $U(n) = O(n)$. We can also define the **special unitary group** of degree n over a field F by $SU(n) = \{A \in SL_n(V) : A^{-1} = A^*\}$.*

Note that for $V = \mathbb{R}$, we have that $U(n)$ is a real vector space and not a complex vector space because multiplying by an i scalar does not preserve the complex conjugate transpose.

Before we move on, we will prove the following theorem, which we will refer to again in the final chapter.

Theorem 2.1.5. *Let A be an operator for a vector space V and $v \in V$ be an eigenvector with a corresponding eigenvalue λ (so $Av = \lambda v$). Then for another operator B of V, if $AB = BA$ and we presume that eigenvalues are distinct (you can't have two different eigenvectors get you the same eigenvalue), we have $Bv = mv$, where m is another eigenvalue corresponding to v.*

Proof. Let $Av = \lambda v$ and suppose that $AB = BA$. Then $ABv = BAv = B\lambda v = \lambda Bv$ (since scalars commute). Since we can't have two different eigenvectors give us the same eigenvalue, Bv must be a scalar multiple of v. Thus, for a scalar m, $Bv = mv$. $\square$

We will now define an inner product and prove a result about a common inner product used.

Definition 2.1.6. *Let V be a complex vector space. An* ***inner product*** *on V is a map $V \times V \mapsto \mathbb{C}$ given by $(u, v) \mapsto \langle u, v\rangle$ with the following properties:*

1. *Linearity in the second factor: $\langle u, v_1 + av_2\rangle = \langle u, v_1\rangle + a\langle u, v_2\rangle$ for all $u, v_1, v_2 \in V$ and $a \in \mathbb{C}$.*

2. *Conjugate symmetry: $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.*

3. *Positivity: For all $v \in V$, the quantity $\langle v, v\rangle \geq 0$, with $\langle v, v\rangle = 0$ if and only if $v = 0$.*

Note that the first two points imply that $\langle v_1 + av_2, u\rangle = \langle v_1, u\rangle + \overline{a}\langle v_2, u\rangle$.

Theorem 2.1.7. *For the vector space $M_n(\mathbb{C})$, the* ***Hilbert-Schmidt inner product*** *is given by the formula $\langle A, B\rangle = tr(A^*B)$ and, as its name suggests, is an inner product.*

Proof. We will show that it meets the three conditions required to be an inner product. Let $A, B, C \in M_n(\mathbb{C})$ and $\alpha \in \mathbb{C}$.

1. Linearity in the second factor: We have $\langle A, B + \alpha C\rangle = tr(A^*(B + \alpha C)) = tr(A^*B + \alpha A^*C) = tr(A^*B) + \alpha tr(A^*C) = \langle A, B\rangle + \alpha\langle A, C\rangle$.

2. Conjugate symmetry: We have $\langle A, B\rangle = tr(A^*B) = tr(A^*(B^*)^*) = tr((B^*A)^*) = tr(\overline{(B^*A)}^T) = tr(\overline{(B^*A)})$, since transpose does not affect the diagonal entries, which is all that matters when taking the trace. This is, by definition, $\overline{\langle B^*, A\rangle}$.

3. Positivity: We have $\langle A, A\rangle = tr(A^*A)$. Note that when you multiply matrices, the diagonal entries of the new matrix product are simply the diagonal entries of the previous matrices multiplied together. Since A^* is the conjugate transpose, and transposing does not affect diagonal entries, we are essentially multiplying the diagonal entries of $\overline{A}$ with A, and adding them up since we are taking that trace. But multiplying a number by its conjugate results in squares being added together, so the trace must be greater than or equal to 0. The sum can only be 0 if each entry was 0. Thus, the positivity condition is satisfied.

Thus, this is indeed an inner product. □

Recall that a unitary operator U has the property that $\langle Uv, Uw\rangle = \langle v, w\rangle$ for all v, w. Note that $\langle Uv, Uw\rangle = UU^*\langle v, w\rangle$ since $\langle Uv, w\rangle = U\langle v, w\rangle$ and $\langle v, Uw\rangle = U^*\langle v, w\rangle$ by rules of an inner product. Also recall that a unitary matrix, say U, has the property that $U^{-1} = U^*$, or UU^* is the identity. These both being called unitary makes sense since $\langle Uv, Uw\rangle = \langle v, w\rangle \iff UU^*\langle v, w\rangle = \langle v, w\rangle \iff UU^* = 1 \iff U^* = U^{-1}$.

Now we will go over the concept of normalization, which will be useful in our last chapter.

Definition 2.1.8. *Let V be a finite-dimensional vector space, $\langle \cdot, \cdot\rangle$ be an inner product on V, and $v \in V$. If $\langle v, v\rangle = 1$, then we say that v is* ***normalized****. We say that a vector v can be normalized if multiplying it by a real number results in a normalized vector.*

Theorem 2.1.9. *Let V be a finite-dimensional vector space with an inner product defined by $\langle v, v\rangle = v^*v$ for all $v \in V$. If $v \neq 0$, then v can be normalized.*

Proof. We defined our inner product as $v^*v = \langle v, v\rangle$, which we know (by definition of inner product) must be greater than 0 for a nonzero vector v. So we have that $\langle v, v\rangle = a > 0$, where $a \in \mathbb{R}$. If we let $w = \frac{v}{\sqrt{a}}$, then $\langle w, w\rangle = \langle \frac{v}{\sqrt{a}}, \frac{v}{\sqrt{a}}\rangle = \frac{1}{\sqrt{a}}\frac{1}{\sqrt{a}}\langle v, v\rangle$

by rules of inner product (where $(\frac{1}{\sqrt{a}})^* = \frac{1}{\sqrt{a}}$ since $a \in \mathbb{R}$). This then equals $\frac{1}{a}a = 1$.Thus, w is normalized, and so we have shown that any nonzero vector v can be normalized. □

Now we will go over a more complicated group and a theorem utilizing it.

Definition 2.1.10. *Let*

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

The ***complex symplectic group*** *is defined by* $Sp_n(\mathbb{C}) = \{A \in GL_{2n}(\mathbb{C}) : -\Omega A^T \Omega = A^{-1}\}$. *The* ***real symplectic group*** *is defined by* $Sp_n(\mathbb{R}) = \{A \in GL_{2n}(\mathbb{R}) : -\Omega A^T \Omega = A^{-1}\}$.

Note $-\Omega A^T \Omega = A^{-1} \implies -\Omega A^T \Omega A = I_{2n}$. Also note that $-\Omega = \Omega^{-1}$ and $\Omega^2 = -\Omega$, as seen below:

$$\Omega * (-\Omega) = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ 0 & I_n \end{pmatrix} = I_{2n}, \text{ so } -\Omega = \Omega^{-1}$$

$$\Omega^2 = \Omega\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} = -\Omega.$$

Now that we have discussed the complex symplectic group, we can discuss the following result.

Theorem 2.1.11. $Sp_1(\mathbb{C}) = SL_2(\mathbb{C})$ *and* $Sp_1(\mathbb{R}) = SL_2(\mathbb{R})$

Proof. Note that if n=1, then for an arbitrary matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Sp_n(\mathbb{C})$, we have the following:

$$-\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Computing this tells us that it is true if and only if $ad-bc=1$. This means that the determinant must be 1 for all $A \in Sp_1(\mathbb{C})$. Thus, we can say that $Sp_1(\mathbb{C}) = SL_2(\mathbb{C})$. It follows, using the same process, that $Sp_1(\mathbb{R}) = SL_2(\mathbb{R})$. □

2.2 Dual Space

In this section, we go over some of the basics of the dual space; namely, its definition and two resulting theorems.

Definition 2.2.1. *For a complex vector space V, the* ***dual space*** *of V is defined by $V^* = Hom(V, \mathbb{C}) = \{f : V \to \mathbb{C} : f$ is a linear transformation$\}$.*

Theorem 2.2.2. *Let V be a complex vector space. Then its dual space V^* is a complex vector space.*

Proof. Since for all $f \in Hom(V, \mathbb{C})$, f is (by definition) a linear transformation, we have that closure under addition, associativity under addition, commutativity under addition, scalar distributivity, and scalar associativity all hold.

Now, if $f_0 : V \to \{0\}$, then f_0 is a linear transformation from V to $\mathbb{C}$, and f_0 is the additive identity of $Hom(V, \mathbb{C})$. Also, since V has an identity element for scalar multiplication, call it 1, we also have it for any linear transformation in $Hom(V, \mathbb{C})$. This also gives us that every linear transformation has an additive inverse, as we just multiply any f by -1. Thus, $Hom(V, \mathbb{C})$ is a complex vector space. □

Theorem 2.2.3. *Let V be a complex vector space with basis $\{e_1 \ldots e_n\}$. Then its dual space V^* has a basis $\{f_1, ..., f_n\}$, where $f : V \to \mathbb{C}$ and $f_i(e_j) = 1$ if $i = j$ and $f_i(e_j) = 0$ if $i \neq j$. Note that this theorem implies that $dim(V) = dim(V^*)$.*

Proof. We want to show that $\{f_1, \ldots, f_n\}$ is a basis for V^*. So we need to show that it is linearly independent and spans V^*. We will start by showing linear independence.

Let a_i be a scalar for all $1 \leq i \leq n$. Suppose $a_1 f_1 + \cdots + a_n f_n = 0$. Then $(a_1 f_1 + \cdots + a_n f_n)(e_i) = 0$, for some $1 \leq i \leq n$. Since f_i is a linear transformation for all i, we have $a_1 f_1(e_i) + \cdots + a_n f_n(e_i) = 0$. But $f_i(e_j) \neq 0$ if $i \neq j$, so we get that $a_i f_i(e_i) = 0$. But since $f_i(e_j) = 1$ if $i = j$, we get that $a_i = 0$. Since a_i was chosen arbitrarily, this must be true for all i. Thus, we have linear independence.

Now we want to show that our potential basis spans V^*. Let $f \in V^*$. We want to show that $f = f(e_1)f_1 + \cdots + f(e_n)f_n$. Note that $(f(e_1)f_1 + \cdots + f(e_n)f_n)(e_i) = f(e_1)f_1(e_i) + \cdots + f(e_n)f_n(e_i) = f(e_i)$ (by similar reasoning as above). Thus, we can say that $f = f(e_1)f_1 + \cdots + f(e_n)f_n$. □

2.3 Tensor Products

In this section, we discuss the complicated topic of tensor products. Since these are difficult to understand for those who have never studied them, we will try to give some intuition and lean gently into the topic before formally defining them. Much of the beginning of this section comes from a YouTube video created by Jim Fowler (a mathematics professor at Ohio State University), known simply as *Tensor products* [3]. It helped the author understand the big idea behind tensor products, and so we are now introducing tensor products in the same manner here.

Let U, V, and W be finite-dimensional vector spaces. Then we can take the bilinear space $U \times V$ and map it into W, or $U \times V \to W$. Note that this is a bilinear map since $U \times V$ is bilinear. The idea of tensor products is to turn this bilinear map into a linear map, which is easier to understand.

Say $U \otimes V$ is the vector space of all possible linear combinations of $u \otimes v$, where $u \in U$, $v \in V$, and the following conditions are satisfied:

1. $u \otimes v_1 + u \otimes v_2 = u \otimes (v_1 + v_2)$ for all $u \in U$ and $v_1, v_2 \in V$

2. $u_1 \otimes v + u_2 \otimes v = (u_1 + u_2) \otimes v$ for all $u_1, u_2 \in U$ and $v \in V$

3. $(\lambda u) \otimes v = \lambda(u \otimes v) = u \otimes (\lambda v)$ for all $u \in U$, $v \in V$, and scalars λ.

We set up our conditions this way so that the linear map $U \otimes V \to W$ is the same as the bilinear map $U \times V \to W$. (Again, we want to turn our original bilinear map into an easier to work with linear map.)

For example, for a bilinear map $f : U \times V \to W$, we have $f(u, v_1 + v_2) = f(u, v_1) + f(u, v_2)$. For the corresponding linear map $g : U \otimes V \to W$, we have $g(u \otimes (v_1 + v_2)) = g(u \otimes v_1 + u \otimes v_2) = g(u \otimes v_1) + g(u \otimes v_2)$. Note the similarity between these two maps in how they function, but the tensor product map (which we will rigorously define in a little) is linear. Now let's look at a specific example.

Example 2.3.1. *Note that the vector space $\mathbb{R}^2$ has dimension 2. Say $\{e_1, e_2\}$ is a basis for $\mathbb{R}^2$. Then the tensor product $\mathbb{R}^2 \otimes \mathbb{R}^2$ is spanned by $\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$. If we wanted to add, for example, the first two basis elements, we would have $(e_1 \otimes e_1) + (e_1 \otimes e_2) = e_1 \otimes (e_1 + e_2)$. Note that we can not simplify, for example, $(e_1 \otimes e_1) + (e_2 \otimes e_2)$, as everything in a vector space is a sum of its basis elements (and a tensor product is a vector space).*

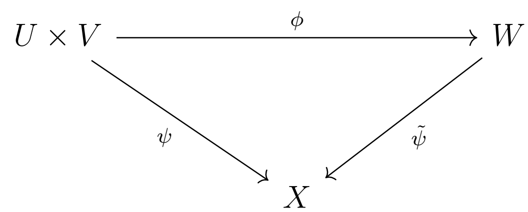
Also, note that $(e_1 + e_2) \otimes (e_1 + e_2) = (e_1 \otimes e_1) + (e_1 \otimes e_2) + (e_2 \otimes e_1) + (e_2 \otimes e_2)$, which is the sum of our basis elements. If we had a bilinear map f, then $f(a + b, a + b) = f(a, a) + f(a, b) + f(b, a) + f(b, b)$, and so again we see how our linear tensor product map functions like a bilinear map.

We can also see in this example that $dim(\mathbb{R}^2 \otimes \mathbb{R}^2) = 4$, and it is a fact that $dim(U \otimes V) = dim(U)dim(V)$ for all finite-dimensional vector spaces U, V.

Now we that we have the general idea behind tensor products, let us formally define it.

Definition 2.3.2. *If U and V are finite-dimensional real or complex vector spaces, then a tensor product of U with V is a vector space W, together with a bilinear map $\phi : U \times V \to W$ with the following property: if ψ is any bilinear map of $U \times V$*

into a vector space X, there exists a unique linear map $\tilde{\psi}$ of W into X such that the following diagram commutes:



Note that the bilinear map ψ from $U \times V$ into X turns into the linear map $\tilde{\psi}$ of W into X.

There are many fascinating results within the area of tensor products, but we will only mention the following (although we will return to them when we discuss tensor products of representations).

Theorem 2.3.3. *Let U, V be finite-dimensional vector spaces. Then:*

1. *The tensor product $U \otimes V$ exists*

2. *$dim(U \otimes V) = dim(U)dim(V)$.*

3. *For $A : U \to U$ and $B : V \to V$, where A and B are linear operators, there exists a unique linear operator from $U \otimes V$ to $U \otimes V$, denoted $A \otimes B$, such that $(A \otimes B)(u \otimes v) = (Au) \otimes (Bv)$ for all $u \in U$ and $v \in V$. Furthermore, if A_1, A_2 are linear operators on U and B_1, B_2 are linear operators on V, then $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1A_2) \otimes (B_1B_2)$.*

We omit the proofs of these theorems, although they can be found in Chapter 4 of Hall's book [4]. For further reading on tensor products, see Chapter 10 of Bruce N. Cooperstein's *Advanced linear algebra* (2015) [2].

2.4 Topology

In this section, we go over some topological background relevant to our discussion of Lie theory. This is not meant to be a detailed discussion of all topology relevant to the topic, but instead a brief introduction so that we can use this information later. The definitions, examples, and results contained here are considered common knowledge for the subject and serve as a useful introduction for those unfamiliar with the subject. We begin by defining a topological space.

Definition 2.4.1. *A* ***topological space*** *is a pair* (X, τ) *of sets where* τ *is a collection of subsets of* X *satisfying the following:*

1. *Both* $\emptyset \in \tau$ *and* $X \in \tau$.

2. *If* $\{X_i\}_{i \in I} \subseteq \tau$*, then* $\cup_{i \in I} X_i \in \tau$*, where* I *is a set that can be infinitely uncountable.*

3. *If* $X_1, \ldots, X_n \in \tau$*, then* $X_1 \cap \cdots \cap X_n \in \tau$.

The sets in τ *are called the open sets.*

Note that a topological space is not necessarily a vector space. We will now move on to examples of topological spaces.

Example 2.4.2. *These first two are examples of topological spaces.*

- $X = \{1, 2, 3\}, \tau = \{\emptyset, X\}$

- $X = \{1, 2, 3\}, \tau = \{\emptyset, X, \{1\}, \{2\}, \{1, 2\}\}$

The following is NOT a topology: $X = \{1, 2, 3\}$ *with* $\tau = \{\emptyset, X, \{1, 2\}, \{2, 3\}\}$*. This is not a topology since the intersection of the last two subsets is* $\{2\}$ *and that is not in* τ.

Now we will define what it means for a subset to be closed and then go over an example and theorem.

Definition 2.4.3. *Let (X, τ) be a topological space. Then a subset C of X is* ***closed*** *if $X \setminus C$ is open (which means it is in τ). So the complement of an open set is closed and the complement of a closed set is open.*

Example 2.4.4. *Let $X = \{1, 2, 3\}$ and let $\tau = \{\emptyset, X, \{1\}, \{2\}, \{1, 2\}\}$. Since $X \setminus \{3\} \in \tau$, that set is open. Thus, $\{3\} = X \setminus \{1, 2\}$ is closed. Similarly, $\emptyset$ is closed since $X \setminus \emptyset = X \in \tau$ and is therefore open. Finally, X is closed since $X \setminus X = \emptyset \in \tau$ and is therefore open.*

Theorem 2.4.5. *If $f : X \to Y$ is any function, then $f^{-1}(Y \setminus C) = X \setminus f^{-1}(C)$.*

Proof. Note that $x \in f^{-1}(Y \setminus C) \iff f(x) \in Y \setminus C \iff f(x) \notin C \iff x \notin f^{-1}(C) \iff x \in X \setminus f^{-1}(C)$. Thus, we have equality. □

Now we define what it means for a function between topological spaces to be continuous and prove a result using this.

Definition 2.4.6. *Let X and Y be topological spaces. A function $f : X \to Y$ is continuous if for all $U \subseteq Y$ open, $f^{-1}(U)$ is open in X.*

Theorem 2.4.7. *Say $f : X \to Y$ is continuous. If $C \subseteq Y$ is closed, then $f^{-1}(C) \subseteq X$ is closed.*

Proof. Let $f : X \to Y$ be a continuous function and let $C \subseteq Y$ be closed. So $Y \setminus C$ is open, which implies that $f^{-1}(Y \setminus C)$ is open since f is continuous. By the previous theorem, this is equivalent to $X \setminus f^{-1}(C)$, so this must also be open. Thus, $f^{-1}(C)$ is closed because the complement of an open set is closed. □

We now note the following before moving forward.

Definition 2.4.8. *Any set of a single element is known as a* ***singleton****.*

Theorem 2.4.9. *A singleton is closed in a topological space.*

It is also a fact that Lie groups, which we will discuss in the next chapter, are topological spaces. So they are topological groups. We omit the details of this. We can now prove the next theorem.

Theorem 2.4.10. *The kernel of a Lie group homomorphism is closed.*

Proof. Let $\Phi : G \to H$ be a Lie group homomorphism. Then $ker(\Phi) = \{X \in G : \Phi(X) = e_H\} = \Phi^{-1}(\{e_H\})$. Since all singletons are closed, $\{e_H\}$ is closed. Thus, by a previous theorem, $\Phi^{-1}(\{e_H\}) = ker(\Phi)$ is closed. □

Our final theorem in this section will be useful later, but we mention it here as it is relevant to topology.

Theorem 2.4.11. *Let V and W be finite-dimensional real or complex vector spaces. Then any linear transformation $T : V \to W$ is continuous. More importantly, we note that a continuous function commutes with a limit.*

Chapter 3

Introduction to Lie Theory

This section will cover Lie Groups, Lie Algebras, their connection, and various results stemming from them.

3.1 Lie Groups and the Exponential

Definition 3.1.1. *A* ***matrix Lie group*** *is any subgroup G of $GL_n(\mathbb{C})$ such that if A_m is any sequence of matrices in G and A_m converges to some matrix A, then either $A \in G$ or $A \notin GL_n(\mathbb{C})$ (which would mean A is not invertible). We also call G a* ***closed subgroup*** *of $GL_n(\mathbb{C})$.*

Examples of matrix Lie groups include the following:

- The general linear groups, $GL_n(\mathbb{C})$ and $GL_n(\mathbb{R})$, since they are subgroups of the general linear group
- The special linear groups, $SL_n(\mathbb{C})$ and $SL_n(\mathbb{R})$ (the determinant is a continuous function, so if A_m is a sequence of matrices with determinant 1 and A_m converges to A, then A also has determinant 1).
- The orthogonal and special orthogonal groups, $O(n)$ and $SO(n)$.

- The unitary and special unitary groups, $U(n)$ and $SU(n)$.
- The complex and real symplectic groups mentioned earlier, $Sp_n(\mathbb{C})$ and $Sp_n(\mathbb{R})$.

An important topic when discussing Lie groups is the exponential, whose definition we show below.

Definition 3.1.2. *If X is an $n \times n$ matrix, we define the* ***exponential of X****, denoted e^X, by the (usual) power series $e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}$, where X^0 is defined to be the identity matrix I_n and where X^m is the repeated matrix product of X with itself.*

It is a fact that e^X converges for all $X \in M_n(\mathbb{C})$ and that e^X is a continuous function of X. We also have various properties of e^X that will be useful for us.

Theorem 3.1.3. *Let X and Y be arbitrary $n \times n$ matrices. Then the following hold:*

1. $e^0 = I_n$
2. $(e^X)^* = e^{X^*}$
3. e^X *is invertible and* $(e^X)^{-1} = e^{-X}$
4. $e^{(a+b)X} = e^{aX} e^{bX}$ *for all* $a, b \in \mathbb{C}$
5. *If* $XY = YX$*, then* $e^{X+Y} = e^X e^Y = e^Y e^X$
6. *If* $A \in GL_n(\mathbb{C})$*, then* $e^{AXA^{-1}} = Ae^X A^{-1}$
7. $e^{(X,Y)} = (e^X, e^Y)$.

We omit the proofs of the first 6 results, although they can be found in Chapter 2 of Hall's book [4]. We will, however, prove the final result.

Proof. We have $e^{(X,Y)} = \sum_{n=0}^{\infty} \frac{(X,Y)^n}{n!}$, by definition. To raise an ordered pair to a power n, you simply raise each component to n, and so we get $\sum_{n=0}^{\infty} \frac{(X^n,Y^n)}{n!}$. Since this is a summation of an ordered pair over $n!$, we can "distribute" the summation and

$n!$ into each component, yielding $(\sum_{n=0}^{\infty} \frac{X^n}{n!}, \sum_{n=0}^{\infty} \frac{Y^n}{n!})$. This is then, by definition, (e^X, e^Y). □

The next theorem is useful as well.

Theorem 3.1.4. *Let X be an $n \times n$ complex matrix and $t \in \mathbb{R}$. Then $e^{tX} \in M_n(\mathbb{C})$ and $\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X$. In particular, when evaluating at $t = 0$, we get X.*

Note that in general, the **derivative** (as we will call this) of e^{X+tY} is not equal to Ye^{X+tY}. We also note that general derivative rules work, including the product and chain rules, which we will utilize.

We will now go over a few interesting results before diving into the next section, on Lie algebras.

Theorem 3.1.5. *Let T be a linear transformation between finite-dimensional real or complex vector spaces. Then the derivative function commutes with $T(e^{tX})$.*

Proof. Say $T : M_n(\mathbb{C}) \to M_n(\mathbb{C})$. (This same process will work when utilizing real entries.) Then $\frac{d}{dt}T(e^{tX}) = \lim_{h\to 0} \frac{T(e^{(t+h)X}) - T(e^{tX})}{h}$ (note that the derivative is with respect to t). Since T is a linear transformation, we can pull the T out to get $\lim_{h\to 0} T(\frac{e^{(t+h)X} - e^{tX}}{h})$. Finally, since T is continuous, it will commute with this limit. But this new limit is $\frac{d}{dt}(e^{tX})$. So we have $T\frac{d}{dt}(e^{tX})$. Thus, we have commutativity. □

Theorem 3.1.6. *Every invertible $n \times n$ matrix can be expressed as e^X for some $X \in M_n(\mathbb{C})$. Since Lie groups are subsets of invertible matrices, all elements of a Lie group can be expressed as e^X for some $X \in M_n(\mathbb{C})$.*

Theorem 3.1.7. *For any $X \in M_n(\mathbb{C})$, $det(e^X) = e^{tr(X)}$.*

Proof. If X is diagonalizable with eigenvalues λ_i, then e^X is diagonalizable with eigenvalues e^{λ_i} (note that diagonalizing a matrix does not change its eigenvalues). Thus, the trace of X is the sum of its eigenvalues and $det(e^X) = e^{\lambda_1}...e^{\lambda_n}$ (since the determinant of a matrix is the product of its eigenvalues), and this equals $e^{\lambda_1 \cdots \lambda_n}$ (by our properties of exponentials since eigenvalues commute), and this finally equals $e^{tr(X)}$ (since $tr(X)$ is the sum of its eigenvalues when diagonlized). We omit the proof of the case where X is not diagonalizable. □

Theorem 3.1.8. *Let*

$$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix},$$

where $a, b, c \in \mathbb{C}$. Then $e^{tA} \in SL_2(\mathbb{C})$ for all $t \in \mathbb{R}$.

Proof. The theorem is really asking us to show that $\det(e^{tA}) = 1$ for all $t \in \mathbb{R}$. Note that A is a matrix with complex entries. Since $t \in \mathbb{R}$, we know that the matrix tA will also be a matrix of the same form as A. So we can essentially ignore the t and just say tA is equivalent to a matrix of the same form as A, call it X.

Note $e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}$. Let's look at a simpler case of X, where $b = c = 0$. Then we have

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

Expanding this e^X out, we get

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} + \frac{\begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}^2}{2!} + \text{... and so on.}$$

This means that

$$e^X = \begin{pmatrix} \sum_{m=0}^{\infty} \frac{a^m}{m!} & 0 \\ 0 & \sum_{m=0}^{\infty} \frac{(-a)^m}{m!} \end{pmatrix},$$

which implies that

$$\det(e^X) = e^a e^{-a} = e^0 = 1.$$

Now, how do we generalize this? Note that this specific case is really a diagonal matrix. So it is enough to show that our matrix X (which is basically the same as A) is diagonalizable. If we can show that X is diagonalizable then we are good because that means it can be of the form written above and therefore the determinant is 1!

Let $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$. Let's find some eigenvalues:

$$|A - \lambda I| = (a-\lambda)(-a-\lambda) - bc = 0 \implies -a^2 + \lambda^2 - bc = 0 \implies \lambda = \pm\sqrt{a^2+bc}.$$

Let λ equal the positive square root and $-\lambda$ equal the negative square root.

Now we find some eigenvectors:

$$\begin{pmatrix} a-\lambda & b \\ c & -a-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Then we get $ax - \lambda x + by = 0 \implies x = \frac{-by}{a-\lambda}$. If we do the same thing for $-\lambda$, we get $x = \frac{-by}{a+\lambda}$. Letting $y = 1$ gives us the eigenvector matrix as follows:

$$P = \begin{pmatrix} \frac{-b}{a-\lambda} & \frac{-b}{a+\lambda} \\ 1 & 1 \end{pmatrix}.$$

So $|P| = \frac{-b}{a-\lambda} + \frac{b}{a+\lambda} = \frac{-2b\lambda}{a^2-\lambda^2}$. Now we know how to find the inverse of a 2x2 matrix; you take our original matrix P, then you switch the positions of the main diagonal, multiply the off-diagonal by-1, and divide the whole thing by $|P|$. So we get that

$$P^{-1} = \begin{pmatrix} \frac{\lambda^2-a^2}{2b\lambda} & \frac{\lambda-a}{2\lambda} \\ \frac{a^2-\lambda^2}{2b\lambda} & \frac{a+\lambda}{2\lambda} \end{pmatrix}.$$

To actually diagonalize this matrix, we compute $P^{-1}AP$.

The rest just involves computation. We get for our final diagonalized matrix the following:

$$D = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}.$$

Thus, we have shown that X is diagonalizable, and therefore e^X has determinant 1. □

3.2 Lie Algebras

Our main topic for this thesis is Lie theory and, in particular, Lie algebras. We will define it using a more linear algebra approach first, where we define it as a specific type of vector space, and then we will define it with its relation to Lie groups. And we will show that these two definitions are actually equivalent!

Definition 3.2.1. *A finite-dimensional real or complex* ***Lie algebra*** *is a finite-dimensional real or complex vector space* $\mathfrak{g}$ *together with a map* $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *that satisfies the following properties:*

1. $[\cdot,\cdot]$ *is bilinear, or* $[X + \lambda Y, Z] = [X, Z] + \lambda[Y, Z]$ *and* $[X, Y + \lambda Z] = [X, Y] + \lambda[X, Z]$ *for all* $X, Y, Z \in \mathfrak{g}$ *and* λ *a scalar in our chosen field.*

2. $[\cdot,\cdot]$ *is skew-symmetric, or* $[X, Y] = -[Y, X]$ *for all* $X, Y \in \mathfrak{g}$. *Note that this implies* $[X, X] = 0$ *for all* $X \in \mathfrak{g}$.

3. *The Jacobi identity holds, or* $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ *for all* $X, Y, Z \in \mathfrak{g}$.

We say that two elements X and Y of a Lie algebra $\mathfrak{g}$ **commute** if $[X, Y] = 0$, and we say that a Lie algebra is **commutative** if this is true for all of its elements. The map $[\cdot,\cdot]$ is referred to as the **bracket operation** on $\mathfrak{g}$. Also, note that Lie algebras are algebras (which may not have associativity), where the “vector multiplication” is defined by our Lie bracket.

We will now prove some examples of vector spaces being Lie algebras.

Theorem 3.2.2. *Let $[\cdot,\cdot] : \mathbb{R}^3 \times \mathbb{R}^3$ be given by $[x, y] = x \times y$, where $x \times y$ is the cross product. Then $\mathbb{R}^3$ is a Lie algebra.*

Proof. Note that $\mathbb{R}^3$ is a vector space, so we just need to show that the cross product is bilinear, skew-symmetric, and satisfies the Jacobi identity. Let $(x_1, x_2, x_3) = x, (y_1, y_2, y_3) = y, (z_1, z_2, z_3) = z \in \mathbb{R}^3$ and $\lambda \in \mathbb{R}$. Recall that the cross product $x \times y$ is defined by $x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$.

1. Bilinearity: We have $[x + \lambda y, z] = (x + \lambda y) \times z$
$= ((x_2+\lambda y_2)z_3-(x_3+\lambda y_3)z_2, (x_3+\lambda y_3)z_1-(x_1+\lambda y_1)z_3, (x_1+\lambda y_1)z_2-(x_2+\lambda y_2)z_1)$
$= (x_2z_3 + \lambda y_2z_3 - x_3z_2 + \lambda y_3z_2, x_3z_1 + \lambda y_3z_1 - x_1z_3 - \lambda y_1z_3, x_1z_2 + \lambda y_1z_2 - x_2z_1 - \lambda y_2z_1)$
$= (x_2z_3 - x_3z_2, x_3z_1 - x_1z_3, x_1z_2 - x_2z_1) + (\lambda y_2z_3 + \lambda y_3z_2, \lambda y_3z_1 - \lambda y_1z_3, \lambda y_1z_2 - \lambda y_2z_1)$
$= (x \times z) + (\lambda y \times z) = [x, z] + \lambda[y, z]$. Similarly, we have that $[x, y + \lambda z] = [x, y] + \lambda[x, z]$. So we have bilinearity.

2. Skew-symmetry: Note that $[x, y] = x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) = -(-x_2y_3 + x_3y_2, -x_3y_1 + x_1y_3, -x_1y_2 + x_2y_1) = -(y_2x_3 - y_3x_2, y_3x_1 - y_1x_3, y_1x_2 - y_2x_1) = -(y \times x) = -[y, x]$. So we have skew-symmetry.

3. Jacobi identity: To show that the Jacobi identity is satisfied, we need to show that $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$. Note that $[x, [y, z]] = [x, y \times z] = x \times (y \times z)$
$= x \times (y_2z_3 - y_3z_2, y_3z_1 - y_1z_3, y_1z_2 - y_2z_1)$. Looking at only the first column, we end up with $x_2(y_1z_2 - y_2z_1) - x_3(y_3z_1 - y_1z_3) = x_2y_1z_2 - x_2y_2z_1 - x_3y_3z_1 + x_3y_1z_3$. Similarly, the first column of $[y, [z, x]]$ is $y_2z_1x_2 - y_2z_2x_1 - y_3z_3x_1 + y_3z_1x_3$ and the first column of $[z, [x, y]]$ is $z_2x_1y_2 - z_2x_2y_1 - z_3x_3y_1 + z_3x_1y_3$. Adding these

three columns together does indeed result in 0, and the same follows if we were to compute this for the second and third columns. Thus, the Jacobi identity is satisfied.

Thus, $\mathbb{R}^3$ is a Lie algebra. □

Theorem 3.2.3. *Let $\mathcal{A}$ be an associative algebra and let $\mathfrak{g}$ be a subspace of $\mathcal{A}$ such that $XY - YX \in \mathfrak{g}$ for all $X, Y \in \mathfrak{g}$. Then $\mathfrak{g}$ is a Lie algebra with the bracket operation given by the commutator, or $[X,Y] = XY - YX$, where $X, Y \in \mathfrak{g}$.*

Proof. Since $\mathfrak{g}$ is a subspace, and therefore a vector space, we just need to show that the commutator is bilinear, skew-symmetric, and satisfies the Jacobi identity. Let $X, Y, Z \in \mathfrak{g}$ and λ be a scalar in our chosen field.

1. Bilinearity: We have $[X + \lambda Y, Z] = (X + \lambda Y)Z - Z(X + \lambda Y)$
   $= XZ + \lambda YZ - ZX + Z\lambda Y = XZ - ZX + \lambda YZ - Z\lambda Y = [X, Z] + \lambda[Y, Z]$.
   Similarly, we have the $[X, Y + \lambda Z] = [X, Y] + \lambda[X, Z]$. So we have bilinearity.

2. Skew-symmetric: Note that $[X, Y] = XY - YX = -(YX - XY) = -[Y, X]$.
   So we have skew-symmetry.

3. Jacobi identity: Note that $[X, [Y, Z]] = [X, YZ - ZY] = X(YZ - ZY) - (YZ - ZY)X$
   $= XYZ - XZY - YZX + ZYX$. Similarly, $[Y, [Z, X]] = YZX - YXZ - ZXY + XZY$ and $[Z, [X, Y]] = ZXY - ZYX - XYZ + YXZ$. Adding these three results together gives us 0, and so the Jacobi identity is satisfied.

Thus, $\mathfrak{g}$ is a Lie algebra under the commutator. □

Now, note the following definition.

Definition 3.2.4. *Define $\mathfrak{sl}_n(\mathbb{C})$ by $\mathfrak{sl}_n(\mathbb{C}) = \{X \in M_n(\mathbb{C}) : tr(X) = 0\}$.*

Based on our last theorem, we have the following.

Example 3.2.5. *We have that* $\mathfrak{sl}_n(\mathbb{C})$ *is a Lie algebra with the bracket operation given by the commutator, or* $[X, Y] = XY - YX$*, where* $X, Y \in \mathfrak{sl}_n(\mathbb{C})$*.*

For the remainder of this thesis, the bracket operation will be given by the commutator.

We will now list various definitions relation to Lie algebras. Many of these have analogous definitions for other algebraic objects (groups, rings, etc.) and so they should seem familiar. We begin with a subalgebra.

Definition 3.2.6. *A* ***subalgebra*** *of a real or complex Lie algebra* $\mathfrak{g}$ *is a subspace* $\mathfrak{h}$ *of* $\mathfrak{g}$ *such that* $[H_1, H_2] \in \mathfrak{h}$ *for all* $H_1, H_2 \in \mathfrak{h}$*. If* $\mathfrak{g}$ *is a complex Lie algebra and* $\mathfrak{h}$ *is a real subspace of* $\mathfrak{g}$ *which is closed under brackets, then* $\mathfrak{h}$ *is said to be a real subalgebra of* $\mathfrak{g}$*.*

Just like how in ring theory, certain subrings are known as ideals, we can define certain subalgebras as ideals.

Definition 3.2.7. *A subalgebra* $\mathfrak{h}$ *of a Lie algebra* $\mathfrak{g}$ *is said to be an* ***ideal in*** $\mathfrak{g}$ *if* $[X, H] \in \mathfrak{h}$ *for all* $X \in \mathfrak{g}$ *and* $H \in \mathfrak{h}$*.*

And just like how we have centers of groups, we have centers of Lie algebras.

Definition 3.2.8. *The* ***center of a Lie algebra*** $\mathfrak{g}$ *is the set of all* $X \in \mathfrak{g}$ *for which* $[X, Y] = 0$ *for all* $Y \in \mathfrak{g}$*.*

An important aspect of algebra is finding relationships between objects. We usually accomplish this goal by defining a structure-preserving map between these two objects, which we call a homomorphism. We also have this for Lie algebras.

Definition 3.2.9. *If* $\mathfrak{g}$ *and* $\mathfrak{h}$ *are Lie algebras, then a linear map* $\phi : \mathfrak{g} \to \mathfrak{h}$ *is called a* ***Lie algebra homomorphism*** *if* $\phi([X, Y]) = [\phi(X), \phi(Y)]$ *for all* $X, Y \in \mathfrak{g}$*.*

Of course, the most interesting homomorphisms are those that show two algebraic objects are essentially the same. In other words, there is an isomorphism between the two objects. We define a Lie algebra isomorphism exactly as one would think.

Definition 3.2.10. *A bijective Lie algebra homomorphism is called a* ***Lie algebra isomorphism****.*

We can also define automorphisms of Lie algebras just like one would expect.

Definition 3.2.11. *A Lie algebra isomorphism of a Lie algebra with itself is called a* ***Lie algebra automorphism****.*

Now that we have talked about maps between Lie algebras, we define the adjoint map, which will be very useful later.

Definition 3.2.12. *If* $\mathfrak{g}$ *is a Lie algebra and* $X \in \mathfrak{g}$*, define the linear map* $ad_X : \mathfrak{g} \to \mathfrak{g}$ *by* $ad_X(Y) = [X, Y]$*. The map* $X \to ad_X$ *is the* ***adjoint map*** *or* ***adjoint representation****.*

Note that using the definition, we can write $[X, [X, [X, [X, Y]]]] = (ad_X)^4(Y)$. We also have a useful theorem following from the adjoint map definition.

Theorem 3.2.13. *The adjoint map is a Lie algebra homomorphism.*

Proof. If $\mathfrak{g}$ is a Lie algebra, then $ad_{[X,Y]} = ad_X ad_Y - ad_Y ad_X = [ad_X, ad_Y]$. So $ad : \mathfrak{g} \to End(\mathfrak{g})$ is a Lie algebra homomorphism. □

Since Lie algebras are, by definition, vector spaces, it makes sense to define some analogous vector space definitions for Lie algebras: we define a direct sum of Lie algebras, a Lie algebra decomposition, an irreducible Lie algebra, and a simple Lie algebra.

Definition 3.2.14. *If* $\mathfrak{j}$ *and* $\mathfrak{k}$ *are Lie algebras, the* ***direct sum*** *of* $\mathfrak{j}$ *and* $\mathfrak{k}$ *is the vector space direct sum of* $\mathfrak{j}$ *and* $\mathfrak{k}$ *with the bracket operation given by* $[(X_1, X_2), (Y_1, Y_2)] = ([X_1, Y_1], [X_2, Y_2])$ *for all* $X_1, Y_1 \in \mathfrak{j}$ *and* $X_2, Y_2 \in \mathfrak{k}$ *.*

Definition 3.2.15. *If* $\mathfrak{g}$ *is a Lie algebra and* $\mathfrak{j}$ *and* $\mathfrak{k}$ *are subalgebras, we say that* $\mathfrak{g}$ ***decomposes*** *as the Lie algebra direct sum of* $\mathfrak{j}$ *and* $\mathfrak{k}$ *if* $\mathfrak{g}$ *is the direct sum of* $\mathfrak{j}$ *and* $\mathfrak{k}$ *as vector spaces with* $[X_1, X_2] = 0$ *for all* $X_1 \in \mathfrak{j}, X_2 \in \mathfrak{k}$.

Definition 3.2.16. *A Lie algebra* $\mathfrak{g}$ *is called* ***irreducible*** *if the only ideals in* $\mathfrak{g}$ *are* $\mathfrak{g}$ *and* $\{0\}$.

Definition 3.2.17. *A Lie algebra is called* ***simple*** *if it is irreducible and* $dim(\mathfrak{g}) \geq 2$.

As an enlightening example, we will prove the following theorem for $Lie(SL_2(\mathbb{C}))$.

Example 3.2.18. *The Lie algebra* $Lie(SL_2(\mathbb{C}))$ *is simple.*

Proof. A basis for $Lie(SL_2(\mathbb{C}))$ is the following:

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \text{ and } Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Recall that the bracket operation for this Lie algebra is defined as a commutator. So $[X, Y] = XY - YX = Z$. Similarly, $[Z, X] = 2X$ and $[Z, Y] = -2Y$.

Now we want to show that this Lie algebra is simple, so we want to show that it is irreducible and has dimension greater than or equal to 2. We have 3 basis elements, so the dimension is 3, which is indeed greater than 2. Now suppose that $\mathfrak{h}$ is a non-trivial ideal of our Lie algebra. Then if W is in our Lie algebra and $H \in \mathfrak{h}$, we have by definition that $[W, H] \in \mathfrak{h}$. Now, since $H \in \mathfrak{h}$, H is in our Lie algebra, so it can be written as a linear combination of our basis elements. Say $H = aX + bY + cZ$, with at least one of the scalars a, b, c non-zero. We want to show that $\mathfrak{h} = \mathfrak{sl}_2(\mathbb{C})$, as that would imply that the Lie algebra is irreducible (and therefore simple since our dimension is 2 or greater).

Suppose that b is the only non-zero scalar. Then $[X, [X, H]] = [X, [-2cX + bZ]] = -2bX$ is a non-zero multiple of X (this can be seen through matrix calculations).

Since $\mathfrak{h}$ is an ideal, $[X, H] \in \mathfrak{h}$, which means that $[X, [X, H]] \in \mathfrak{h}$, which means $-2bX \in \mathfrak{h}$, which means that $X \in \mathfrak{h}$ since we can multiply $-2bX$ by the scalar $-\frac{1}{2b}$ to get X.

Since X is in our ideal, $[Y, X]$ is also in our ideal. But $[Y, X]$ is a nonzero multiple of Z, so Z is also in our ideal. Similarly, $[Y, [Y, X]]$ is in our ideal. Since it is a nonzero multiple of Y, Y is in our ideal. So we have the basis elements X, Y, and Z in our ideal. Since our ideal is a subspace that includes all basis elements of the Lie algebra, it is equal to the Lie algebra.

A similar process works when choosing a being the only non-zero scalar and c being the only non-zero scalar. Since we know each of these cases work individually, a linear combination of these cases will work because our commutator is bilinear. Thus, our ideal is equal to our Lie algebra and we have a simple Lie algebra.

□

3.3 Connecting Lie Groups and Lie Algebras

Now that we have discussed what Lie algebras are, our goal is to make a clear connection between Lie algebras and Lie groups. We will first give an alternative definition of a Lie algebra, where we define it as being associate with a Lie group.

Definition 3.3.1. *If G is a Lie group that is a subgroup of $GL_n(\mathbb{C})$, then its* ***associated Lie algebra*** *is defined by $Lie(G) = \{X \in M_n(\mathbb{C}) : e^{tX} \in G \text{ for all } t \in \mathbb{R}\}$.*

Note that, by definition, $Lie(G)$ is a subset of $M_n(\mathbb{C})$, and we have that $G \neq \mathfrak{g}$ (an immediate reason being that $\mathfrak{g}$ contains the zero matrix, while G does not).

Since it is not clear that $Lie(G)$ is a Lie algebra using our earlier definition, we need to show that. This proof will be less rigorous than some of our other proofs, but the goal is to illustrate the idea in a convincing way more so than rigorously prove it.

Theorem 3.3.2. *If G is a Lie group, then $Lie(G)$ is a Lie algebra (using our first definition) under the commutator bracket.*

Proof. Note that $Lie(G)$ is a subset of $M_n(\mathbb{C})$, or $M_n(\mathbb{R})$, which will use the same proof techniques as below. So we need to show that $Lie(G)$ is a subspace that has a Lie bracket (the commutator in this case).

Let $A, B \in Lie(G)$. Then e^{tA} and e^{tB} are in G for all $t \in \mathbb{R}$. We want to show that $e^{t(A+B)}$ is in G, as that would imply that $A + B \in Lie(G)$, which means that we would have vector addition. The following is where this proof is less rigorous, as we have not covered why the following is true, but for the sake of completeness, we will include it.

We have $e^{t(A+B)} = \lim_{m\to+\infty}(e^{\frac{tA}{m}}e^{\frac{tB}{m}})^m$, where what is inside the limit is also in G. Since G is closed (as it is a Lie group), the limit must also be in G. Thus, $e^{t(A+B)}$ is in G and so $A + B \in Lie(G)$.

Now we want to show that scalar multiplication holds. Let $z \in \mathbb{C}$. We want to show that $zA \in Lie(G)$, or that $e^{tzA} \in G$. Note that if $z \in \mathbb{R}$, this works because tz would become our new scalar in $\mathbb{R}$. We actually don't know this is true for all $z \in \mathbb{C}$, so we end up with the fact that $Lie(G)$ becomes a real Lie algebra, but not necessarily a complex Lie algebra.

Finally, we want to show that the Lie bracket holds under the commutator operation. We will not be showing that each of the properties (bilinearity, skew-symmetric, and the Jacobi identity) holds, as that follows from how we defined our commutator, but we will show that there is closure.

Note that we have $[A, B] = AB - BA$, as defined by our commuator. We want to show that $AB - BA \in Lie(G)$. Note that by a previous theorem, if $A \in G$, then $XAX^{-1} \in Lie(G)$. So we can have $e^{tA}Be^{-tA} \in Lie(G)$. If we evaluate the derivative of $(e^{tA}B)(e^{-tA})$ at $t = 0$ (by utilizing the product rule), we get that $(AB)e^0 + (e^0B)(-A) = AB - BA$, so we have that $AB - BA \in Lie(G)$, which means our commutator holds. Thus, $Lie(G) = Lie(G)$ is indeed a (real) Lie algebra. $\square$

3.4 Examples of Lie Algebras

This section goes over some fun examples of Lie algebras, which help to illuminate some of the concepts we have been talking about.

Theorem 3.4.1. *$Lie(GL_n(\mathbb{C})) = M_n(\mathbb{C})$ and $Lie(GL_n(\mathbb{R})) = M_n(\mathbb{R})$*

Proof. Note that $Lie(GL_n(\mathbb{C})) = \{X \in M_n(\mathbb{C}) : e^{tX} \in GL_n(\mathbb{C}) \text{ for all } t \in \mathbb{R}\}$ by definition. We know, by our properties of the exponential, that e^X is invertible for all $X \in M_n(\mathbb{C})$. Since t is merely a scalar in the real numbers, e^{tX} is also invertible for all $X \in M_n(\mathbb{C})$, or $e^{tX} \in GL_n(\mathbb{C})$ for all $X \in M_n(\mathbb{C})$. Thus, $Lie(GL_n(\mathbb{C}))$ is the set of all elements in $M_n(\mathbb{C})$ i.e., $Lie(GL_n(\mathbb{C})) = M_n(\mathbb{C})$. This works similarly with $Lie(GL_n(\mathbb{R})) = M_n(\mathbb{R})$. $\square$

Note that from this point on, we will not be specifying each case for both the real and complex entries of a matrix. In most cases, they will both work out, so we will only show it is true for the complex entries.

The next few theorems show some equivalent ways of looking at specific Lie algebras.

Theorem 3.4.2. *Note that $Lie(SL_n(\mathbb{C})) = \{X \in M_n(\mathbb{C}) : e^{tX} \in SL_n(\mathbb{C})$ for all $t \in \mathbb{R}\}$ and $\mathfrak{sl}_n(\mathbb{C}) = \{X \in M_n(\mathbb{C}) : tr(X) = 0\}$. Then $Lie(SL_n(\mathbb{C})) = \mathfrak{sl}_n(\mathbb{C})$.*

Proof. Say $A \in Lie(SL_n(\mathbb{C}))$. We want to show that $A \in \{X \in M_n(\mathbb{C}) : tr(X) = 0\}$. Since $A \in M_n(\mathbb{C})$ already, we just need to show that $tr(A) = 0$. Since $A \in Lie(SL_n(\mathbb{C}))$, that means $e^{tA} \in SL_n(\mathbb{C})$, which means that $det(e^{tA}) = 1$. By a previous theorem, we know that $det(e^{tA}) = e^{tr(tA)}$. So $1 = e^{tr(tA)} = e^{t*tr(A)}$, since t is a scalar. In order to have $e^{t*tr(A)} = 1$, we need $t * tr(A)$ to equal 0, $tr(A)$ must equal

0 (because this has to work for all $t \in \mathbb{R}$). Thus, $A \in \mathfrak{sl}_n(\mathbb{C})$ and so $Lie(SL_n(\mathbb{C}))$ is a subset of $\mathfrak{sl}_n(\mathbb{C})$.

Now let $A \in \mathfrak{sl}_n(\mathbb{C})$. Then $tr(A) = 0$. We want to show that $A \in Lie(SL_n(\mathbb{C}))$. Since $A \in M_n(\mathbb{C})$ already, we just need to show that $e^{tA} \in SL_n(\mathbb{C})$ for all $t \in \mathbb{R}$. So we need to show that $det(e^{tA}) = 1$. Again, this means that we want to show that $e^{tr(tA)} = 1$, or that $e^{t*tr(A)} = 1$. Since $tr(A) = 0$, this is true! Thus, $A \in Lie(SL_n(\mathbb{C}))$ and $\mathfrak{sl}_n(\mathbb{C})$ is a subset of $Lie(SL_n(\mathbb{C}))$. Thus, since our sets are subsets of each other, they are equal! □

As an exercise, we will show that $Lie(SL_n(\mathbb{C}))$ is a Lie algebra with the now-equivalent $\mathfrak{sl}_n(\mathbb{C})$ definition. Note, however, that we already know this is true since $Lie(G)$ is always a Lie algebra for a matrix Lie group G. From this point on, we will use the notation $Lie(SL_n(\mathbb{C}))$ to also refer to $\mathfrak{sl}_n(\mathbb{C})$.

Theorem 3.4.3. *$Lie(SL_n(\mathbb{C}))$ is a Lie algebra (using our first definition of a Lie algebra).*

Proof. Note that $Lie(SL_n(\mathbb{C}))$ is a subset of $M_n(\mathbb{C})$, so we only need to show that it is a subspace that also satisfies our bracket operation.

We will first show that we have vector addition. Let $A, B \in Lie(SL_n(\mathbb{C}))$. Then $tr(A) = tr(B) = 0$, so $0 = 0 + 0 = tr(A) + tr(B) = tr(A + B)$ (since adding A and B means you are adding their diagonals, and since the diagonal of each of those add to be 0, the diagonal of their sum will as well). Thus, since $tr(A + B) = 0$, $A + B \in Lie(SL_n(\mathbb{C}))$ and we have vector addition.

Now we will show that scalar multiplication holds.Let $z \in \mathbb{C}$ and $A \in Lie(SL_n(\mathbb{C}))$. Then $tr(zA) = z * tr(A) = z * 0 = 0$. Thus, $zA \in Lie(SL_n(\mathbb{C}))$ and we have scalar multiplication. Thus, we have a subspace. Now we just need to show that there exists a Lie bracket. We will define it as the commutator, or $[A, B] = AB - BA$. We need to show that closure works with the commutator.

Using our same A and B, note that $tr(AB - BA) = tr(AB) - tr(BA)$ and this equals 0 since $tr(AB) = tr(BA)$ (utilizing linear algebra). Thus, $AB - BA \in Lie(SL_n(\mathbb{C}))$. Since $[A, B] = AB - BA$, our commutator satisfies closure. This means that we have a Lie bracket - note that because this Lie bracket also works for $M_n(\mathbb{C})$, all three properties (bilinearity, skew-symmetric, and the Jacobi identity) will also carry over to this subspace. Thus, we have a Lie algebra. $\square$

For the next several theorems, we will show equivalent ways of looking at Lie algebras. We will notice that when taking the associated Lie algebra of a matrix Lie group, we generally change from using determinants to using traces in our sets, as well as changing the adjoint from equaling its multiplicative inverse to equaling its additive inverse. Informally, think of Lie algebras as using additive properties, whereas Lie groups use multiplicative properties.

Theorem 3.4.4. *$Lie(U(n)) = \{X \in M_n(\mathbb{C}) : X^* = -X\}$.*

Proof. Say $X \in M_n(\mathbb{C})$ and that $e^{tX} \in U(n)$ for all $t \in \mathbb{R}$. Then $(e^{tX})^* = (e^{tX})^{-1}$ for all $t \in \mathbb{R}$. By our properties of exponents, this means that $e^{tX^*} = e^{t(-X)}$ for all $t \in \mathbb{R}$. If we use our derivative rule to take the derivative of both sides and evaluate them at $t = 0$, we find that $X^* = -X$. Thus, we have half of the proof complete.

Say $X^* = -X$. We want to show that $e^{tX} \in U(n)$ for all $t \in \mathbb{R}$. Note that $(e^{tX})^* = e^{(tX)^*} = e^{tX^*} = e^{t(-X)} = e^{-tX} = (e^{tX})^{-1}$, by our exponential properties and the fact that $X^* = -X$. So we have containment with both sides and thus our sets are equal! $\square$

Theorem 3.4.5. *$Lie(SU(n)) = \{X \in M_n(\mathbb{C}) : X^* = -X, tr(X) = 0\}$.*

Proof. Let $X \in Lie(SU(n))$. Then $e^{tX} \in SU(n)$, so $(e^{tX})^* = (e^{tX})^{-1}$, which means that $e^{tX^*} = e^{-tX}$. Evaluating the derivative at $t = 0$ on both sides yields $X^* = -X$. Also, we have that $1 = det(e^{tX}) = e^{tr(tX)}$ (by a previous theorem), which equals $e^{t*tr(X)}$. In order for $e^{t*tr(X)}$ to equal 1, the exponent must equal 0, which means that

$t * tr(X) = 0$. Since t can be any scalar from $\mathbb{R}$, we must have $tr(X) = 0$. Thus, we have $X^* = -X$ and $tr(X) = 0$.

Now we will show the other direction. Let $X \in M_n(\mathbb{C})$ such that $-X = X^*$ and $tr(X) = 0$. Note then that $e^{-tX} = e^{tX^*}$, which means that $(e^{tX})^* = (e^{tX})^{-1}$. We also have that $det(e^{tX}) = e^{tr(tX)} = e^{t*tr(X)} = e^0 = 1$. Thus, $e^{tX} \in SU(n)$ and therefore $X \in Lie(SU(n))$. So we have double containment and our sets are equal. □

Theorem 3.4.6. *$Lie(O(n)) = Lie(SO(n)) = \{X \in M_n(\mathbb{C}) : X^T = -X, tr(X) = 0\}$.*

Proof. Let $X \in Lie(O(n))$. Then $e^{tX} \in O(n)$, which means that $e^{tX^T} = e^{-tX}$, which implies that $X^T = -X$ (utilizing our favorite derivative rules). Note that $X^T = -X$ implies $tr(X) = 0$, as taking the transpose of a matrix leaves the diagonals alone, so you have each entry equaling its negative, which must mean they are all 0, which means adding them gives 0, which means the trace is 0. So we have $1 = e^0 = e^{t*tr(X)} = e^{tr(tX)} = det(e^{tX})$. So $e^{tX} \in SO(n)$ and thus $Lie(O(n)) \subseteq Lie(SO(n))$.

Now let $X \in Lie(SO(n))$. Then $e^{tX} \in SO(n)$,which means that $e^{tX^T} = e^{-tX}$, which implies that $X^T = -X$ (as we just said). We also have that $1 = det(e^{tX}) = e^{t*tr(X)}$, which implies that $tr(X) = 0$. So $Lie(SO(n) \subseteq \{X \in M_n(\mathbb{C}) : X^* = -X, tr(X) = 1\}$.

Finally, let $X \in \{X \in M_n(\mathbb{C}) : X^T = -X, tr(X) = 0\}$. So $X^T = -X$, which means $e^{tX^T} = e^{-tX}$. So $X \in Lie(O(n))$. Thus, we have $Lie(O(n)) \subseteq Lie(SO(n)) \subseteq \{X \in M_n(\mathbb{C}) : X^T = -X, tr(X) = 1\} \subseteq Lie(O(n))$, which means they are all equal and we are done. □

Theorem 3.4.7. *$Lie(Sp_n(\mathbb{C})) = \{X \in M_{2n}(\mathbb{C}) : \Omega X^T \Omega = X\}$.*

Proof. Let $X \in Lie(Sp_n(\mathbb{C}))$. Then $e^{tX} \in Sp_n(\mathbb{C})$. Then $\Omega^{-1}(e^{tX})^T \Omega = e^{-tX}$. Note that the left-hand side yields $\Omega^{-1} e^{tX^T} \Omega = e^{t\Omega^{-1} X^T \Omega}$ (by rules of our exponent), which equals $e^{-t\Omega X^T \Omega}$, since $\Omega^{-1} = -\Omega$ (see earlier example). So we have $e^{-t\Omega X^T \Omega} = e^{-tX}$.

If we utilize our derivative rule and evaluate at $t = 0$, we get that $-\Omega X^T \Omega = -X$. Canceling the negative gives us $\Omega X^T \Omega = X$. So we have one direction complete.

Now let $X \in M_{2n}(\mathbb{C})$ such that $\Omega X^T \Omega = X$. Now, $-\Omega(e^{tX})^T \Omega = -\Omega e^{tX^T} \Omega = e^{-t\Omega X^T \Omega} = e^{-tX}$, since $\Omega X^T \Omega = X$. Thus, $X \in Lie(Sp_n(\mathbb{C})$, and we are done! □

Much like with $Lie(SL_n(\mathbb{C}))$, we will show that $Lie(Sp_n(\mathbb{C}))$ is a Lie algebra with this new, equivalent definition (even though we know this is true since the symplectic group is a matrix Lie group).

Theorem 3.4.8. *$Lie(Sp_n(\mathbb{C}))$ is a Lie algebra (using our first definition).*

Proof. Note $Lie(Sp_n(\mathbb{C}))$ is a subset of $M_{2n}(\mathbb{C})$. Like usual, we need to show vector addition, scalar multiplication, and that a Lie bracket (the commutator) holds.

Let A and B be elements of our Lie algebra. Then $A + B = \Omega A^T \Omega + \Omega B^T \Omega$ $= \Omega(A^T + B^T)\Omega = \Omega(A + B)^T \Omega$. Thus, we have vector addition.

Let $z \in \mathbb{C}$. Then $zA = z\Omega A^T \Omega = \Omega z A^T \Omega = \Omega(zA)^T \Omega$. So we have scalar multiplication.

Now, let $[A, B] = AB - BA$. Like usual, we just need to show closure. So we have $[A, B] = AB - BA = \Omega A^T \Omega \Omega B^T \Omega - \Omega B^T \Omega \Omega A^T \Omega = \Omega(-A^T B^T)\Omega + \Omega(B^T A^T)\Omega$, since $\Omega^2 = -\Omega$. This then equals $\Omega(B^T A^T - A^T B^T)\Omega$, which equals $\Omega(AB - BA)^T \Omega$ (by rules of a transpose from linear algebra). Thus, we have closure and therefore a Lie algebra. □

We have (finally) finished going over examples of Lie algebras and their equivalent definitions, and hopefully it has been insightful. The next theorem is key for later.

Theorem 3.4.9. *The following forms a basis for $Lie(SU(2))$:*

$$A = \begin{pmatrix} \frac{i}{2} & 0 \\ 0 & \frac{-i}{2} \end{pmatrix}, B = \begin{pmatrix} 0 & \frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & \frac{-1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.$$

And the following forms a basis for $Lie(SO(3))$:

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, F = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then there exists a Lie algebra isomorphism between $Lie(SU(2))$ and $Lie(SO(3))$.

Proof. Note that by utilizing the last two examples, we see that both of these Lie algebras have dimension 3 (since they each have 3 basis elements), so they are isomorphic as vector spaces. Now, by using the same matrices from those two examples, we can define a map such that A goes to D, B goes to E, and C goes to F. Note also that $[A, B] = C, [B, C] = A, [C, A] = B$ and $[D, E] = F, [E, F] = D, [F, D] = E$. So this mapping preserves the commutator relations. Since our Lie algebras share the same commutator relations (or satisfies the requirements for a Lie algebra homomorphism) and are isomorphic as vector spaces, they are isomorphic as Lie algebras. □

3.5 Lie Group and Lie Algebra Properties

This section covers some important properties that Lie groups and their corresponding Lie algebras have.

Theorem 3.5.1. *Let G be a matrix Lie group with an associated Lie algebra $Lie(G)$. Then $AXA^{-1} \in Lie(G)$ for all $A \in G$ and $X \in Lie(G)$.*

Proof. Note that $e^{t(AXA^{-1})} = Ae^{tX}A^{-1}$ by our exponential properties. But this is in G since $A \in G$ and $e^{tX} \in G$ (because $X \in Lie(G)$). Thus, $e^{t(AXA^{-1})} \in G$, which means that $AXA^{-1} \in Lie(G)$, and we are done. □

Theorem 3.5.2. *Let G and H be matrix Lie groups, with Lie algebras $Lie(G)$ and $Lie(H)$, respectively. Suppose that $\Phi : G \to H$ is a Lie group homomorphism. Then there exists a unique real-linear map $\phi : Lie(G) \to Lie(H)$ such that $\Phi(e^X) = e^{\phi(X)}$ for all $X \in Lie(G)$. The map ϕ has the following additional properties:*

1. $\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A)^{-1}$, *for all* $X \in Lie(G)$ *and* $A \in G$.

2. $\phi([X, Y]) = [\phi(X), \phi(Y)]$, *for all* $X, Y \in Lie(G)$ *(so this map is a Lie algebra homomorphism).*

3. $\phi(X) = \frac{d}{dt}\Phi(e^{tX})|_{t=0}$ *for all* $X \in Lie(G)$.

So every Lie group homomorphism gives rise to a (unique) Lie algebra homomorphism.

Proof. We will not show the fact that this map is unique, however, we will prove the three properties.

1. Note that $e^{t\phi(AXA^{-1})} = \Phi(Ae^{tX}A^{-1})$ by the properties of our exponential and since t is a scalar that can distribute. This then equals $\Phi(A)\Phi(e^{tX})\Phi(A)^{-1}$ since Φ is a homomorphism, which then equals $\Phi(A)e^{\phi(tX)}\Phi(A)^{-1}$, by what we defined as our unique real-linear map, which then equals $e^{t\Phi(A)\phi(X)\Phi(A)^{-1}}$, once again by our properties. By utilizing our derivative theorem and evaluating at $t = 0$ on our first term before the very first equals sign and this last term, we get that $\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A)^{-1}$.

2. Note that $\phi([X, Y]) = \phi(XY - YX) = \phi(\frac{d}{dt}e^{tX}Ye^{-tX}|_{t=0})$, which can be verified utilizing the product rule. We then know that the derivative commutes with a linear transformation, so we can rewrite this as $\frac{d}{dt}\phi(e^{tX}Ye^{-tX})|_{t=0}$, which then equals $\frac{d}{dt}\Phi(e^{tX})\phi(Y)\Phi(e^{-tX})|_{t=0}$ by the first property of this theorem. Finally, we can rewrite this as $\frac{d}{dt}e^{t\phi(x)}\phi(Y)e^{-t\phi(x)}|_{t=0}$ by how we defined our map, and by utilizing the product rule in reverse, this equals $[\phi(X), \phi(Y)]$.

3. Note that $\frac{d}{dt}\Phi(e^{tX})|_{t=0} = \frac{d}{dt}e^{t\phi(x)}|_{t=0}$ by how we defined our linear map, as well as t being a scalar, and this equals $\phi(X)$ by simply evaluating the derivative. So $\phi(X) = \frac{d}{dt}\Phi(e^{tX})|_{t=0}$.

Thus, we have finished the proof. $\square$

Theorem 3.5.3. *Suppose G, H, and K are matrix Lie groups and $\Phi : G \to H$ and $\Psi : H \to K$ are Lie group homomorphisms. Let $\Lambda : G \to K$ be the composition of Φ and Ψ, and let ϕ, ψ, and λ be the associated Lie algebra maps associated to our groups. Then we have $\lambda = \phi \circ \psi$.*

Proof. For any $X \in Lie(G)$, we have $e^{\lambda(tX)} = \Lambda(e^{tX}) = \Phi(\Psi(e^{tX})) = \Phi(e^{t\psi(X)}) = e^{t\phi(\psi(X))}$. Differentiating both sides at $t = 0$ yields $\lambda(X) = \phi(\psi(X))$, and so the functions are equal. $\square$

Theorem 3.5.4. *If $\Phi : G \to H$ is a Lie group homomorphism and $\phi : Lie(G) \to Lie(H)$ is the associated Lie algebra homomorphism, then $ker(\Phi)$ is a closed, normal subgroup of G and the Lie algebra of the kernel is given by $Lie(ker(\Phi)) = ker(\phi)$.*

Proof. We know, from a beginning Abstract Algebra course, that $ker(\Phi)$ is a normal subgroup of G. We also know, from our detour to topology, that the kernel of a Lie group homomorphism is closed. So we have ker(Φ) is a closed, normal subgroup of G. Now want to show that $Lie(ker(\Phi)) = ker(\phi)$.

Recall that all elements of a Lie group can be written in the form e^{tX}. If $e^{tX} \in ker(\Phi)$ for all $t \in \mathbb{R}$, we have $e^{t\phi(X)} = \Phi(e^{tX}) = I$ for all $t \in \mathbb{R}$, where I is the identity. Differentiating both sides at $t = 0$ yields $\phi(X) = 0$, which means that $X \in ker(\phi)$.

Let $X \in ker(\phi)$. Then $\Phi(e^{tX}) = e^{t\phi(X)} = I$ for all $t \in \mathbb{R}$. Thus, $e^{tX} \in ker(\Phi)$ and therefore $X \in Lie(ker(\Phi))$. Thus, we have that $Lie(ker(\Phi)) = ker(\phi)$. $\square$

Definition 3.5.5. *Let G be a matrix Lie group with an associated Lie algebra $Lie(G)$. Then for each $A \in G$, define a map $Ad_A : Lie(G) \to Lie(G)$ by $Ad_A(X) = AXA^{-1}$. We call this the **adjoint map**.*

Theorem 3.5.6. *The **adjoint map** is an invertible linear transformation.*

Proof. Let $X, Y \in Lie(G)$ and λ be a scalar in our chosen field. Then $Ad_A(X+Y) = A(X+Y)A^{-1} = (AX+AY)A^{-1} = AXA^{-1} + AYA^{-1} = Ad_A(X) + Ad_A(Y)$. Also, $\lambda Ad_A(X) = \lambda AXA^{-1} = A\lambda XA^{-1} = Ad_A(\lambda X)$. Thus, the adjoint map is a linear transformation.

Now let $Ad_A(X) = Ad_A(Y)$. Then $AXA^{-1} = AYA^{-1}$. So we get $A^{-1}AXA^{-1}A = A^{-1}AYA^{-1}A \implies X = Y$. Thus, the adjoint map is injective.

Finally, recall that $AXA^{-1} \in Lie(G)$ for all $A \in G$ and for all $X \in Lie(G)$. So $A^{-1}XA \in Lie(G)$. Thus, $Ad_A(A^{-1}XA) = AA^{-1}XAA^{-1} = X$. Since X is an arbitrary element in $Lie(G)$, the adjoint map is surjective. Thus, we have shown that the adjoint map is an invertible linear transformation. □

Theorem 3.5.7. *Let G be the matrix Lie group with the associated Lie algebra $Lie(G)$. Let $GL(Lie(G))$ denote the group of all invertible linear transformations of $Lie(G)$. Then the map $A \rightarrow Ad_A$ is a homomorphism of G into $GL(Lie(G))$. Also, for each $A \in G$, Ad_A is a Lie algebra homomorphism. We omit the proof of this.*

Since $Lie(G)$ is a finite-dimensional real vector space for a matrix Lie group G, it has dimension n, where $n \in \mathbb{N}$. Then $GL(Lie(G))$ is isomorphic to $GL_n(\mathbb{R})$. Thus, we will regard $GL(Lie(G))$ as a matrix Lie group. It is a fact that $Ad : G \rightarrow GL(Lie(G))$ is a Lie group homomorphism. By a previous theorem, we know that there is an associated real linear map $X \rightarrow Ad_X$ from the Lie algebra of G to the Lie algebra of $GL(Lie(G)$, with the property that $e^{Ad_X} = Ad_{e^X}$.

Theorem 3.5.8. *Let G be a matrix Lie group, $Lie(G)$ be its associated Lie algebra, and $Ad : G \rightarrow GL(Lie(G))$ be as in the last theorem. Let $ad : Lie(G) \rightarrow Lie(GL(Lie(G)))$ be the associated Lie algebra map. Then for all $X, Y \in Lie(G)$, $ad_X(Y) = [X, Y]$.*

Proof. By a previous theorem, we know that $ad_X = \frac{d}{dt} Ad_{e^{tX}}$ when $t = 0$. Thus,

$ad_X(Y) = \frac{d}{dt}e^{tX}Ye^{-tX}$ when $t = 0$. Computing this using the product rule, with $e^{tX}Y$ as the first term and e^{-tX} as the second term, results in $Xe^{tX}Ye^{-tX} + e^{tX}Y(-Xe^{-tX})$. Plugging in $t = 0$ gives us $XY - YX$, which is $[X, Y]$. □

Theorem 3.5.9. *For any $X \in M_n(\mathbb{C})$, let $ad_X : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be given by $ad_X Y = [X, Y]$. Then for any $Y \in M_n(\mathbb{C})$, we have $e^X Y E^{-X} = Ad_{e^X}(Y) = e^{ad_X}(Y)$, where $e^{ad_X}(Y) = Y + [X, Y] + \frac{1}{2}[X, [X, Y]] + ...$ (where we note that this comes from the series definition of e^X).*

Before we go any further, we are going to mention the topic of complexification. Since some of these results can work without necessarily having a Lie group corresponding to our Lie algebra, we will use the notation $\mathfrak{g}$ again.

Definition 3.5.10. *If V is a finite-dimensional vector space, then the **complexification** of V, denoted $V_{\mathbb{C}}$, is the space of formal linear combinations $v_1 + iv_2$, with $v_1, v_2 \in V$.*

Theorem 3.5.11. *For a finite-dimensional vector space V, $V_{\mathbb{C}}$ is a complex vector space if we define $i(v_1 + iv_2) = -v_2 + iv_1$ for all $v_1, v_2 \in V$.*

Proof. Note that since V is a vector space, associativity under vector addition, commutativity under vector addition, scalar distributvity, and scalar associativity all hold for $V_{\mathbb{C}}$. Now let $v_1, v_2, v_3, v_4 \in V$. Then $v_1 + iv_2, v_3 + iv_4 \in V_{\mathbb{C}}$. Then $(v_1 + iv_2) + (v_3 + iv_4) = (v_1 + v_3) + i(v_2 + v_4) \in V_{\mathbb{C}}$, and so we have closure under addition. Since $0 \in V$, we have that $0 + i0 = 0$ is the additive identity in $V_{\mathbb{C}}$. Similarly, we know that $-v_1$ and $-v_2$ are in V, and so we have $-v_1 - iv_2 \in V_{\mathbb{C}}$. Then $v_1 + iv_2 - v_1 - iv_2 = 0$, and so every element in $V_{\mathbb{C}}$ has an additive inverse. Since V has an identity element for scalar multiplication, call it 1, we also have it for $V_{\mathbb{C}}$, as $1(v_1 + iv_2) = 1v_1 + 1iv_2 = v_1 + iv_2$. Thus, $V_{\mathbb{C}}$ is a complex vector space. □

Note that V is a real subspace of $V_{\mathbb{C}}$, as any element $v \in V$ can be represented as $v + i0 = v \in V_{\mathbb{C}}$.

Theorem 3.5.12. *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ be its complexification. Then the bracket operation on $\mathfrak{g}$ has a unique extension to $\mathfrak{g}_{\mathbb{C}}$ that makes $\mathfrak{g}_{\mathbb{C}}$ into a complex Lie algebra.*

Proof. The bracket operation on $\mathfrak{g}_{\mathbb{C}}$ must be bilinear, so it has to be given by $[X_1 + iX_2, Y_1+iY_2] = ([X_1, Y_1]-[X_2, Y_2])+i([X_1, Y_2]+[X_2, Y_1])$. Therefore it is unique. Now we have to check that it is bilinear, skew-symmetric, and satisfies the Jacobi identity. Because $\mathfrak{g}$ is a real Lie algebra, our map is real bilinear and skew-symmetric. Note that being skew-symmetric means that if this is complex linear in the first factor, it is complex linear in the second factor. Thus, we only need to show that $[i(X_1+iX_2), Y_1+iY_2] = i[X_1 + iX_2, Y_1 + iY_2]$ implies $(-[X_2, Y_1] - [X_1, Y_2]) + i([X_1, Y_1] - [X_2, Y_2])$ on both sides. For the Jacobi identity, there is a lot of computation involved.

We have $[X_1+iX_2, [Y_1+iY_2, Z_1+iZ_2]] = [X_1+iX_2, [Y_1, Z_1]-[Y_2, Z_2]+i([Y_1, Z_2]+[Y_2, Z_1])] = [X_1 + iX_2, Y_1Z_1 - Z_1Y_1 - Y_2Z_2 + Z_2Y_2 + iY_1Z_2 - iZ_2Y_1 + iY_2Z_1 - iZ_1Y_2]$. Using bilinearity, we get a total of 32 terms (and it is quite an ugly mess, hence why we are excluding it from this proof). After that step, we can then make X into Y, Y into Z, and Z into X to get 32 new terms. Then we rotate those terms again to get 32 more terms. Everything ends up canceling, which means the Jacobi identity is indeed satisfied. Thus, we are done! □

Theorem 3.5.13. *Suppose that $\mathfrak{g} \subseteq M_n(\mathbb{C})$ is a real Lie algebra and that for all nonzero X in $\mathfrak{g}$, the element iX is not in $\mathfrak{g}$. Then the "abstract" complexification $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g}$ is isomorphic to the set of matrices in $M_n(\mathbb{C})$ that can be expressed in the form $X + iY$, with $X, Y \in \mathfrak{g}$.*

Proof. Consider the map from $\mathfrak{g}_{\mathbb{C}}$ into $M_n(\mathbb{C})$ sending the formal linear combinations $X + iY$ to the linear combination $X + iY$ of matrices. This map is a complex Lie algebra homomorphism. If $\mathfrak{g}$ satisfies the assumption in the statement of the theorem, this map is also injective and thus an isomorphism of $\mathfrak{g}_{\mathbb{C}}$ with $\mathfrak{g} + i\mathfrak{g} \subseteq M_n(\mathbb{C})$. □

With this theorem in hand, we now have the following isomorphisms (unless otherwise indicated, presume the entries are from the complex field):

- $Lie(GL_n(\mathbb{R}))_{\mathbb{C}} \cong Lie(GL_n(\mathbb{C}))$
- $Lie(U(n))_{\mathbb{C}} \cong Lie(GL_n(\mathbb{C}))$
- $Lie(SU(n))_{\mathbb{C}} \cong Lie(SL_n(\mathbb{C}))$
- $Lie(SL_n(\mathbb{R}))_{\mathbb{C}} \cong Lie(SL_n(\mathbb{C}))$
- $Lie(SO(n))_{\mathbb{C}} \cong Lie(SO_n(\mathbb{C}))$
- $Lie(Sp_n(\mathbb{R}))_{\mathbb{C}} \cong Lie(Sp_n(\mathbb{C})$
- $Lie(Sp_n(\mathbb{C})_{\mathbb{C}} \cong Lie(Sp_n(\mathbb{C})$

For a detailed example of one of these, see the following theorem:

Theorem 3.5.14. $Lie(U(n))_{\mathbb{C}} = Lie(GL_n(\mathbb{C}))$

Proof. We will do proof by containment. Note that $Lie(GL_n(\mathbb{C})) = M_n(\mathbb{C})$. So any element in $Lie(U(n))_{\mathbb{C}}$ must be in $Lie(GL_n(\mathbb{C}))$. Thus, we have one side of containment done.

Now, say $X \in Lie(GL_n(\mathbb{C})) = M_n(\mathbb{C})$. Note that $X = \frac{X-X^*}{2} + i\frac{X+X^*}{2i}$. Now, $(\frac{X-X^*}{2})^* = \frac{X^*-X}{2} = -(\frac{X-X^*}{2})$. Thus, since the complex transpose equals the negative, this is in $Lie(U(n))$. Similarly, $(\frac{X+X^*}{2i})^* = \frac{X^*+X}{-2i} = -(\frac{X+X^*}{2i})$, which would mean it is also in $Lie(U(n))$. So by closure, the sum of these elements must also be in $Lie(U(n))$, which means that multiplying the second element by i and adding it to the first element results in something in $Lie(U(n))_{\mathbb{C}}$. But this new element would be our X. Thus, we have containment in the other direction and so our Lie algebras are equal. □

Note that Lie algebras can have isomorphic complexifications without the original Lie algebras being isomorphic. For example, although both $Lie(SU(2))_{\mathbb{C}}$ and $Lie(SL_2(\mathbb{R}))_{\mathbb{C}}$ are isomorphic to $Lie(SU(2))_{\mathbb{C}}$, the Lie algebra $Lie(SU(2))$ is not isomorphic to the Lie algebra $Lie(SL_2(\mathbb{R}))$.

Theorem 3.5.15. *Let $\mathfrak{g}$ be a real Lie algebra, $\mathfrak{g}_{\mathbb{C}}$ be its complexification, and $\mathfrak{h}$ and arbitrary complex Lie algebra. Then every real Lie algebra homomorphism of $\mathfrak{g}$ into $\mathfrak{h}$ extends uniquely to a complex Lie algebra homomorphism of $\mathfrak{g}_{\mathbb{C}}$ into $\mathfrak{h}$. This is known as the* ***Universal Property of Complexification of Real Lie Algebras****.*

Proof. The unique extension is given by $\pi(X + iY) = \pi(X) + i\pi(Y)$ where $X, Y \in \mathfrak{g}$. This map is a homomorphism of complex Lie algebras, so we are done. □

In the next chapter, we will use our now vast knowledge of Lie theory to combine it with the fascinating subject of representation theory!

Chapter 4

Representation Theory

This chapter will go over the mixing of Lie theory with representation theory. Weights and roots are also covered and will be mentioned again in the final chapter centered on physics.

4.1 Basics

Representation theory is about representing algebraic structures as linear transformations of vector spaces. It is common to look at mappings that go from a group to some matrix, which itself can be represented as a linear transformation. Even though group theory is fairly well-studied, sometimes it is easier to look at groups through the lens of matrices and linear transformations. The formal definition of a representation of a group is defined as follows:

Definition 4.1.1. *A* ***representation of a group*** G *is a homomorphism* $\pi : G \mapsto GL(V)$ *for some finite-dimensional complex vector space* V*.*

More information on the representation theory of finite groups can be found in Benjamin Steinberg's *Representation theory of finite groups* (2012) [6]. We will be focusing on representations involving Lie theory and presume the reader has no background knowledge of the subject.

4.2 Representations of Lie Groups and Algebras

Note that for a finite-dimensional vector space V, $Lie(GL(V)) = End(V)$, where $End(V)$ is the vector space of linear transformations from V into V. These all have dimension n^2.

We have already defined a representation of a finite group, and the definition of a representation of a Lie group is very similar.

Definition 4.2.1. *Let G be a matrix Lie group. A* ***finite-dimensional complex representation*** *of G is a Lie group homomorphism $\Pi : G \to GL(V)$, where V is a finite-dimensional complex vector space (with dimension greater than 0). If V is a finite-dimensional real vector space, then we call Π a* ***finite-dimensional real representation****.*

The definition of a Lie algebra representation is also very similar.

Definition 4.2.2. *Let $\mathfrak{g}$ be a Lie algebra. A* ***finite-dimensional complex representation*** *of $\mathfrak{g}$ is a Lie algebra homomorphism $\pi : \mathfrak{g} \to M_n(\mathbb{V})$, where V is a finite-dimensional complex vector space (with dimension greater than 0). If V is a finite-dimensional real vector space, then we call π is a* ***finite-dimensional real representation****.*

We also have a specific name for an injective representation.

Definition 4.2.3. *A* ***faithful*** *representation is a representation that is an injective homomorphism.*

Note that if $\Pi : G \to GL(V)$ is faithful, then $\{\Pi(A) : A \in G\} \cong G$.

Now we can return to our definition of irreducible from earlier and apply it in this new context, using some new definitions.

Definition 4.2.4. *Let* Π *be a finite-dimensional real or complex representation of a matrix Lie group* G *acting on a space* V*. A subspace* W *of* V *is called* ***invariant*** *if* $\Pi(A)w \in W$ *for all* $w \in W$ *and for all* $A \in G$*.*

Definition 4.2.5. *An invariant subspace* W *is called* ***nontrivial*** *if* $W \neq \{0\}$ *and* $W \neq V$*.*

Definition 4.2.6. *A representation with no nontrivial invariant subspaces is called* ***irreducible****.*

A relatively easy example of an irreducible representation is the trivial representation, of either a Lie group or Lie algebra, as these are mapped into $\mathbb{C}$, which has no nontrivial subspaces. This is because $dim(\mathbb{C}) = 1$ and you can't have a subspace with a lower dimension. Since there are no nontrivial subspaces, there must also be no nontrivial invariant subspaces.

We will now define a specific linear map between representations, which we can then use to create an isomorphism definition for representations.

Definition 4.2.7. *Let* G *be a matrix Lie group,* Π *be a representation of* G *acting on the space* V*, and* Σ *be a representation of* G *acting on the space* W*. A linear map* $\phi : V \to W$ *is called an* ***intertwining map*** *of representations if* $\phi(\Pi(A)v) = \Sigma(A)\phi(v)$ *for all* $A \in G$ *and for all* $v \in V$*. There is an analogous definition for intertwining maps of representations of Lie algebras.*

The following diagram helps illustrate the previous definition:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ \Pi(A) \downarrow & & \downarrow \Sigma(A) \\ V & \xrightarrow{\phi} & W \end{array}$$

Definition 4.2.8. *If* ϕ *is an intertwining map of representations and* ϕ *is invertible* *(or bijective), then ϕ is said to be an **isomorphism** of representations. We use the standard isomorphism notation $\cong$ to denote isomorphic representations.*

We now arrive at our first theorem of representation theory. We will omit the proof, as it follows similarly from previous work we have done.

Theorem 4.2.9. *Let G be a matrix Lie group with corresponding Lie algebra $Lie(G)$ and let Π be a finite-dimensional real or complex representation of G acting on the space V. Then there is a unique representation π of $Lie(G)$ acting on the same space such that $\Pi(e^X) = e^{\pi(X)}$ for all $X \in Lie(G)$. The representation π can be computed as $\pi(X) = \frac{d}{dt}\Pi(e^{tX})$ when $t = 0$. This satisfies $\pi(AXA^{-1}) = \Pi(A)\pi(X)\Pi(A)^{-1}$ for all $X \in Lie(G)$ and for all $A \in G$.*

We note that not every representation π of $Lie(G)$ comes from Π of G, although we are not proving this; it is merely a fun fact. Note that we use the term **connected** in the following theorem, although we omit its definition because it is analytic/topological in nature. Instead, it is more important to note that we already looked at examples of connected Lie groups, such as $SL_n(\mathbb{C})$, and so it is applicable. More information can be found in Hall's book [4].

Theorem 4.2.10. *If G is a connected matrix Lie group with Lie algebra $\mathfrak{g}$, then every element $A \in G$ can be written in the form $A = e^{X_1}...e^{X_m}$ for some $X_1, ..., X_m$ in $\mathfrak{g}$.*

This result allows us to prove some cool results about Lie groups representations and their associated Lie algebra representations. If the Lie group is connected, then irreducibility of one implies irreducibility of the other, and the same can be said for isomorphisms! We formalize this in the next few theorems.

Theorem 4.2.11. *Let G be a connected matrix Lie group with corresponding Lie algebra $Lie(G)$. Let Π be a representation of G and π be the associated representation of $Lie(G)$. Then Π is irreducible $\iff$ π is irreducible.*

Proof. Suppose that Π is irreducible. Now let W be a subspace of V that is invariant under $\pi(X)$ for all $X \in Lie(G)$. We want to show that W is either $\{0\}$ or V. Let $A \in G$. Since G is connected, by a previous theorem we know that A can be written as $A = e^{X_1}...e^{X_m}$ for some $X_1, ..., X_m$ in $Lie(G)$. Since W is invariant under $\pi(X_j)$, it will also be invariant under $e^{\pi(X_j)} = I + \pi(X_j) + \frac{\pi(X_j)^2}{2} + ...$ (from the series definition of the exponential function); this is because the sum of invariant representations is also invariant, and $\pi(X)w \in W$ implies that $\pi^n(X)w \in W$. Thus, we have $\Pi(A) = \Pi(e^{X_1}...e^{X_m}) = \Pi(e^{X_1})...\Pi(e^{X_m}) = e^{\pi(X_1)}...e^{\pi(X_m)}$. Since Π is irreducible and W is invariant under each $\Pi(A)$, W must either be $\{0\}$ or V. Thus, π is irreducible.

Now suppose that π is irreducible and that W is an invariant subspace for Π. Then W is invariant under $\Pi(e^{tX})$ for all $X \in Lie(G)$. Hence, it is also invariant under $\pi(X) = \frac{d}{dt}\Pi(e^{tX})$ evaluated at $t = 0$, as taking the derivative does not mess with invariance. Thus, since π is irreducible, W is either $\{0\}$ or V, and so Π is irreducible. So we're done! □

We will need the following result before moving on.

Theorem 4.2.12. *Let π_1 and π_2 be Lie algebra representations. If $\pi_1 \cong \pi_2$, then $T(e^{\pi_1(X)}v) = e^{\pi_2(X)}T(v)$.*

Proof. Say $\pi_1 \cong \pi_2$. Then there exists a bijective linear transformation $T : V \to W$ such that $T(\pi_1(X)v) = \pi_2(X)T(v)$. Now, $T(e^{\pi_1(X)}(v)) = T((1 + \pi_1(X) + \frac{\pi_1^2(X)}{2} + ...)v) = T(1(v) + \pi_1(X)(v) + ...)$, where we note that everything in the big parenthesis is in V. Since T is a linear transformation, this equals $T(1)(v) + T(\pi_1)(X)(v) + ...$, and so on. But since $\pi_1 \cong \pi_2$, we get $1(T)(v) + \pi_2(X)T(v) + ...$, and so on. Finally, we can factor out a $T(v)$ to get $(1 + \pi_2(X) + ...)T(v)$, which equals $e^{\pi_2(X)}T(v)$, and we're done! □

Theorem 4.2.13. *Let G be a connected matrix Lie group, Π_1 and Π_2 be representations of G, and π_1 and π_2 be the associated Lie algebra representations. Then*

$\Pi_1 \cong \Pi_2 \iff \pi_1 \cong \pi_2$.

Proof. Say $\Pi_1 \cong \Pi_2$. Then there exists a bijective linear transformation $T : V \to W$ such that $T(\Pi_1(A)v) = \Pi_2(A)T(v)$. Then we get the following (where we evaluate at $t = 0$): $T(\pi_1(X)v) = T(\frac{d}{dt}\Pi_1(e^{tX}v) = \frac{d}{dt}T(\Pi_1(e^{tX})v) = \frac{d}{dt}\Pi_2(e^{tX})T(v) = \pi_2(X)T(v)$. Thus, $\pi_1 \cong \pi_2$.

Say $\pi_1 \cong \pi_2$. Then there exists a bijective linear transformation $T : V \to W$ such that $T(\pi_1(X)v) = \pi_2(X)T(v)$. Now, $T(\Pi_1(A)v) = T(\Pi_1(e^{X_1}...e^{X_m})v) = T((\Pi_1(e^{X_1})...\Pi_1(e^{X_m}))v) = T((e^{\pi_1(X)}...e^{\pi_1(X_m)})v)$. Now, evaluating v at all $e^{\pi_1(X_j)}$ yields $T(e^{\pi_1(X_1)}w)$, where $w = (e^{\pi_1(X_2)}...e^{\pi_1(X_m)})v \in V$. But $T(e^{\pi_1(X_1)}w) = e^{\pi_2(X_1)}T(w)$ by our previous theorem. Repeating this process eventually gives us $(e^{\pi_2(X_1)}...e^{\pi_2(X_m)})T(v)$, which equals $\Pi_2(A)T(v)$, and we're done! □

Now that we have these interesting results, we will return to complexification, which we will extend to our new representation theory with the following theorem.

Theorem 4.2.14. *Let $\mathfrak{g}$ be a real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ be its complexification. Then every finite-dimensional complex representation π of $\mathfrak{g}$ has a unique extension to a complex-linear representation of $\mathfrak{g}_{\mathbb{C}}$, also denoted π. Also, π is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$ if and only if it is irreducible as a representation of $\mathfrak{g}$. The extension is given by $\pi(X + iY) = \pi(X) + i\pi(Y)$ for all $X, Y \in \mathfrak{g}$.*

Proof. The existence and uniqueness of the extension follow from an earlier result. Note that a complex subspace W of V is invariant under $\pi(X + iY)$, where $X, Y \in \mathfrak{g}$, if and only if it is invariant under $\pi(X)$ and $\pi(Y)$. Thus, the representation of $\mathfrak{g}$ and its extension have the same invariant subspaces, and so the claim of irreducibility holds. □

Now we apply the concept of unitary for representation theory with the following definition and theorem.

Definition 4.2.15. *If V is a finite-dimensional inner product space and G is a matrix Lie group, a representation $\Pi : G \to GL(V)$ is* ***unitary*** *if $\Pi(A)$ is a unitary operator on V for every $A \in G$.*

Theorem 4.2.16. *Suppose G is a matrix Lie group with corresponding Lie algebra Lie(G). Suppose V is a finite-dimensional inner product space, Π is a representation of G acting on V, and π is the associated representation of Lie(G). If Π is unitary, then $\pi(X)$ is skew self-adjoint for all $X \in$ Lie(G). Conversely, if G is connected, and $\pi(X)$ is skew self-adjoint for all $X \in$ Lie(G), then Π is unitary. Note that we are saying a representation π of a Lie algebra Lie(G) acting on a finite-dimensional inner product space is unitary if $\pi(X)$ is skew self-adjoint for all $X \in$ Lie(G), or $\pi(X)^* = -\pi(X)$ for all $X \in$ Lie(G). (This is consistent with our past work showing that inverse equations in a Lie group get turned into negative equations in a Lie algebra.)*

Proof. Suppose that Π is unitary. Then for all $X \in$ *Lie*(G), we have $(e^{t\pi(X)})^* = \Pi(e^{tX})^* = \Pi(e^{tX})^{-1} = e^{-t\pi(X)}$, where $t \in \mathbb{R}$. Differentiating the leftmost and rightmost sides with respect to t at $t = 0$ reveals that $\pi(X)^* = -\pi(X)$. Now suppose that $\pi(X)^* = -\pi(X)$, then the previous calculation shows that $\Pi(e^{tX}) = e^{t\pi(X)}$ is unitary. But since G is connected, every element in G can be written as a product of exponential, which reveals that $\Pi(A)$ is unitary. Thus, both directions are satisfied and we are done. □

We also have specific types of representations that we can define, namely the standard, trivial, and adjoint representations.

Definition 4.2.17. *Since a matrix Lie group G is a subset of $GL_n(\mathbb{C})$, the map from G into $GL_n(\mathbb{C})$ defined by $\Pi(A) = A$ is a representation of G, and we call this the* ***standard representation*** *of G. This works similarly for a Lie algebra $\mathfrak{g}$, where the map defined by $\pi(X) = X$ is the* ***standard representation*** *of $\mathfrak{g}$.*

Definition 4.2.18. *For any matrix Lie group G, we can define the* ***trivial representation*** $\Pi : G \to GL_1(\mathbb{C}) = \mathbb{C}$ *by* $\Pi(A) = 1$ *for all* $A \in G$*. Since* $\mathbb{C}$ *has no nontrivial subspaces (and therefore no nontrivial invariant subspaces), this is an irreducible representation (as we said earlier). This works similarly for a Lie algebra* $\mathfrak{g}$*, where we have the map* $\pi : \mathfrak{g} \mapsto M_1(\mathbb{C}) = \mathbb{C}$ *defined by* $\pi(X) = 0$ *for all* $X \in \mathfrak{g}$*.*

Definition 4.2.19. *If G is a matrix Lie group with corresponding Lie algebra $Lie(G)$, the* ***adjoint representation*** *of G is the map $Ad : G \to GL(Lie(G))$ given by $A \to Ad_A$. Similarly, the adjoint representation of a finite-dimensional Lie algebra $\mathfrak{g}$ is the map $ad : \mathfrak{g} \to GL(\mathfrak{g})$ given by $X \to ad_X$. Note that by a previous theorem, the Lie algebra representation associated to the adjoint representation of G is in fact the adjoint representation of $\mathfrak{g}$.*

This covers some of the essentials of the representation theory of Lie theory, and so we can move on to an interesting example.

4.3 Representations of Homogeneous Polynomials

Let V_m denote the space of homogeneous polynomials (polynomials whose nonzero terms have the same degree) of degree m in two complex variables. For each $U \in SU(2)$, $\Pi_m(U)$ is defined by $[\Pi_m(U)f](z) = f(U^{-1}z)$, where $z \in \mathbb{C}^2$. Note that on the left-hand side, f is inside the parenthesis, which means that that function is inputting z-values. So its domain is $\mathbb{C}^2$. Similarly, on the right-hand side you will also have in the domain values of $\mathbb{C}^2$, as U^{-1} is a 2×2 matrix and when you multiply that by a 2×1 matrix z, you get a 2×1 matrix. So we write the left-hand side as such so it is clear what our domain is, but what's mainly important is that it is defined by the right-hand side, which looks a lot nicer. We will show that Π_m is a representation of $SU(2)$.

The first step is to show that $\Pi_m(U)$ is a map from V_m to itself. Now, elements of V_m have the form $f(z_1, z_2) = a_0 z_1^m + a_1 z_1^{m-1} z_2 + \cdots + a_m z_2^m$, with $z_1, z_2 \in \mathbb{C}$ and a_j constant in $\mathbb{C}$ for all $j \in \mathbb{Z}$. Note that it is written in this way because each term has to have degree m, as the polynomials are homogeneous. Also note that there are $m+1$ terms, and so $dim(V_m) = m+1$.

Say

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

where $U_{ij} \in SU(2)$, i corresponds to the row of the matrix, and j corresponds to the column of the matrix. Then

$$U^{-1} = \begin{pmatrix} U_{22} & -U_{12} \\ -U_{21} & U_{11} \end{pmatrix}.$$

Then we have

$$[\Pi_m(U)f](z) = [\Pi_m(U)f] \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = f(U^{-1} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix})$$

$$= f(\begin{pmatrix} U_{22} & -U_{12} \\ -U_{21} & U_{11} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}) = f(\begin{pmatrix} U_{22}z_1 - U_{12}z_2 \\ -U_{21}z_1 + U_{11}z_2 \end{pmatrix}).$$

Now, note that U_{22} is equal to U_{11}^{-1}, which is the 11 entry of U^{-1}. Utilizing this logic, we then have the following:

$$f(\begin{pmatrix} U_{11}^{-1}z_1 + U_{12}^{-1}z_2 \\ U_{21}z_1 + U_{22}^{-1}z_2 \end{pmatrix}) = f(U_{11}^{-1}z_1 + U_{12}^{-1}z_2, U_{21}^{-1}z_1 + U_{22}^{-1}z_2)$$

$= a_0(U_{11}^{-1}z_1 + U_{12}^{-1}z_2)^m + a_1(U_{11}^{-1}z_1 + U_{12}^{-1}z_2)^{m-1}(U_{21}^{-1}z_1 + U_{22}^{-1}z_2) + \cdots + a_m(U_{21}^{-1}z_1 + U_{22}^{-1}z_2)^m = \sum_{k=0}^{m} a_k(U_{11}^{-1}z_1 + U_{12}^{-1}z_2)^{m-k}(U_{21}^{-1}z_1 + U_{22}^{-1}z_2)^k$, which has degree m. So

$$\Pi_m(U) : V_m \to V_m.$$

Now we will show that Π_m is a homomorphism. We have $[\Pi_m(U_1)[\Pi_m(U_2)f]](z) = [\Pi_m(U_2)f](U^{-1}z) = f(U_2^{-1}U_1^{-1}z) = f((U_1U_2)^{-1}z) = [\Pi_m(U_1U_2)f](z)$. So we have a homomorphism!

Since Π_m is a map from V_m to V_m and is also a homomorphism, it is a representation!

The associated representation π_m of $Lie(SU(2))$ can be computed (by a previous theorem) as

$$(\pi_m(X)f)(z) = \tfrac{d}{dt}f(e^{-tX}z), \text{ evaluated at } t = 0.$$

Now let $z(t) = (z_1(t), z_2(t))$ be the curve in $\mathbb{C}^2$ defined as $\hat{z}(t) = e^{-tX}$. By the chain rule,

$$\pi_m(X)f = \tfrac{d}{dt}[f(e^{-tX}z)]_{t=0} = \tfrac{d}{dt}[f(z_1(t), z_2(t))]_{t=0} = [\tfrac{\partial f}{\partial z_1}\tfrac{dz_1}{dt}]_{t=0} + [\tfrac{\partial f}{\partial z_2}\tfrac{dz_2}{dt}]_{t=0}.$$

Note that $\frac{d\hat{z}}{dt}$ evaluated at $t = 0$ is the derivative of $e^{-tX}z$ evaluated at $t = 0$, which is $-Xz$. So $\pi_m(X)f = \frac{-\partial f}{\partial z_1}(X_{11}z_1 + X_{12}z_2) - \frac{\partial f}{\partial z_2}(X_{21}z_1 + X_{22}z_2)$, where we get this from recalling that X is a 2×2 matrix and z is a 2×1 matrix.

Now we will take the unique complex linear extension of π to $Lie(SL_2(\mathbb{C})) \cong Lie(SU(2))_{\mathbb{C}}$, as in a previous result.

Now let H, X, and Y be the basis elements of $Lie(SL_2(\mathbb{C}))$:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then $\pi_m(H) = -z_1\frac{\partial}{\partial z_1} + z_2\frac{\partial}{\partial z_2}$, $\pi_m(X) = -z_2\frac{\partial}{\partial z_1}$, and $\pi_m(Y) = -z_1\partial\partial z_2$.

So we have

$$\begin{aligned}\pi_m(H)(z_1^{m-k}z_2^k) = (-z_1\tfrac{\partial}{\partial z_1} + z_2\tfrac{\partial}{\partial z_2})(z_1^{m-k}z_2^k) &= -(m-k)z_1^{m-k}z_2^k + kz_1^{m-k}z_2^k \\ &= (-m+2k)z_1^{m-k}z_2^k.\end{aligned}$$

We also have

$\pi_m(X)(z_1^{m-k}z_2^k) = -z_2\frac{\partial}{\partial z_1}(z_1^{m-k}z_2^k) = -z_2((m-k)z_1^{m-k-1}z_2^k) = -(m-k)z_1^{m-k-1}z_2^{k+1}$.

Finally, we have $\pi_m(Y)(z_1^{m-k}z_2^k) = -z_1\partial\partial z_2(z_1^{m-k}z_2^k) = -z_1(z_1^{m-k}kz_2^{k-1}) = -kz_1^{m-k+1}z_2^{k-1}$.

Recall that an eigenvector v of a matrix A satisfies $Av = \lambda v$ for some scalar λ, which we call the eigenvalue. So $z_1^{m-k}z_2^k$ is an eigenvector for $\pi_m(H)$, with an eigenvalue of $-m+2k$. Since k is arbitrary and $-(m-k)$ is a scalar, $-(m-k)z_1^{m-(k+1)}z_2^{k+1} = -(m-k)z_1^{m-k-1}z_2^{k+1}$ is an eigenvector for $\pi_m(X)$, call it $\hat{v}$. Then $\pi_m(X)\hat{v} = (-m+2(k+1))\hat{v} = (-m+2k+2)\hat{v}$. So the eigenvalue increased by 2. Using $\pi_m(Y)$, we get something similar, except the eigenvalue decreases by 2. The idea of increasing or decreasing an eigenvalue by an integer is very important and will appear later when we discuss roots and weights, as well as when we open our discussion on physics applications. For now, however, we present a theorem:

Theorem 4.3.1. *For $m \geq 0$, the representation π_m is irreducible.*

Proof. We want to show that every nonzero invariant subspace of V_m is equal to V_m. Let W be such a space and let $0 \neq w \in W$. Then $w = a_0z_1^m + a_1z_1^{m-1}z_2 + \cdots + a_mz_2^m$, with at least one $a_k \neq 0$. Let k_0 be the smallest value of k for which $a_k \neq 0$ and consider $\pi_m(X)^{m-k_0}w$. Since each application of $\pi_m(X)$ raises the power of z_2 by 1 (and lowers the power of z_1 by 1), $\pi_m(X)^{m-k_0}$ will kill all terms in w except the $a_{k_0}z_1^{m-k_0}z_2^{k_0}$ term. We show this below.

We have

$$\pi_m(X)(z_1^{m-k}z_2^k) = -(m-k)z_1^{m-k-1}z_2^{k+1}.$$

We also have

$$w = a_{k_0}z_1^{m-k_0}z_2^{k_0} + a_{k_0+1}z_1^{m-k_0-1}z_2^{k_0+1} + \cdots + a_mz_1^0z_2^m.$$

So

$$\pi_m^{m-k_0}(X)(w) = \pi_m^{m-k_0-1}(X)(a_{k_0}(-(m-k_0))z_1^{m-k_0-1}z_2^{k_0+1} + \cdots + a_m(-0)z_1^{-1}z_2^{m+1}).$$

Note that the last term, which is the z_2^m term (when it is in its original form in w, before any derivations) dies off. So doing this $m - k_0$ times will kill $m - k_0$ terms. Since there are $m - k_0 + 1$ terms (as there are m terms after k_0, and then we add 1 to include the k_0 term), we are left with $\pi_m(X)(a_{k_0} z_1^{m-k_0} z_2^{k_0})$.

On the other hand, since $\pi_m(X)(z_1^{m-k} z_2^k) = 0 \implies m = k$, $\pi_m(X)^{m-k_0}$ is a nonzero multiple of z_2^m (since we raised our $z_2^{k_0}$ by degree $(m - k_0)$ to get z_2^m). Since W is invariant, W must contain this multiple of z_2^m, and thus z_2^m. Now, for $0 \leq k \leq m$, we can see that $\pi_m(Y)^k z_2^m$ is a nonzero multiple of $z_1^k z_2^{m-k}$, as $\pi_m(Y) z_2^m = -z_1 \frac{\partial z_2^m}{\partial z_2} = -m z_1 z_2^{m-1} \implies \pi_m(Y)^k z_2^m = \lambda z_1^k z_2^{m-k}$, where λ is a scalar. So W contains $z_1^k z_2^{m-k}$ for all $0 \leq k \leq m$. Since these elements form a basis for V_m, $W = V_m$, and we are done. $\square$

Thus, we have shown that the representations π_m of $Lie(SU(2))$ are irreducible, which we know also implies that the corresponding Lie group representations Π_m of $SU(2)$ are irreducible. The Lie algebra of $SU(2)$ will be discussed again in our section on roots and weights, as well as our chapter on physics applications.

4.4 Tensor Products of Representations

This section will involve taking tensor products, which we presented in our background material chapter, of representations, which we now know a great deal about. However, before we delve into that subject, we present a definition on direct sums of representations, as it is a (relatively) easy-to-understand definition that does not require its own section.

Definition 4.4.1. *Let G be a matrix Lie group and $\Pi_1, \ldots, \Pi_m$ be representations of G acting on the vector spaces $V_1, \ldots, V_m$. Then the* ***direct sum*** *of $\Pi_1, \ldots, \Pi_m$ is a representation $\Pi_1 \oplus \cdots \oplus \Pi_m$ of G acting on the space $V_1 \oplus \cdots \oplus V_m$ defined by* $[\Pi_1 \oplus \cdots \oplus$

$\Pi_m(A)](v_1, \ldots, v_m) = (\Pi_1(A)v_1, \ldots, \Pi_m(A)v_m)$ *for all* $A \in G$*. This works similarly for the Lie algebra side of things, and both of these are indeed representations.*

Now we will define a tensor product of representations.

Definition 4.4.2. *Let* G *and* H *be matrix Lie groups,* Π_1 *be a representation of* G *acting on a space* U*, and* Π_2 *be a representation of* H *acting on a space* V*. Then the* ***tensor product*** *of* Π_1 *and* Π_2 *is a representation* $\Pi_1 \otimes \Pi_2$ *of* $G \times H$ *acting on* $U \otimes V$*, defined by* $(\Pi_1 \otimes \Pi_2)(A, B) = \Pi_1(A) \otimes \Pi_2(B)$ *for all* $A \in G$ *and* $B \in H$*.*

Note that $\Pi_1 \otimes \Pi_2$ in the above definition is indeed a representation. We can also define an analogous definition for Lie algebras and get an interesting theorem.

Theorem 4.4.3. *Let* G *and* H *be matrix Lie groups with Lie algebras Lie*(G) *and Lie*(H)*, respectively, and let* Π_1 *and* Π_2 *be representations of* G *and* H*, respectively. Consider the representation* $\Pi_1 \otimes \Pi_2$ *of* $G \times H$*. If* $\pi_1 \otimes \pi_2$ *denotes the associated representation of Lie*$(G) \oplus$ *Lie*(H)*, then* $(\pi_1 \otimes \pi_2)(X, Y) = \pi_1(X) \otimes I + I \otimes \pi_2(Y)$ *for all* $x \in$ *Lie*(G) *and* $Y \in$ *Lie*(H)*.*

We omit the proof of this theorem, as it requires knowledge of *smooth curves*, although the proof can be found here [4]. We can use this theorem to moitivate the following definition:

Definition 4.4.4. *Let* $\mathfrak{g}$ *and* $\mathfrak{h}$ *be Lie algebras, and let* π_1 *and* π_2 *be their respective representations acting on spaces* U *and* V*. Then the* ***tensor product*** *of* π_1 *and* π_2*, denoted* $\pi_1 \otimes \pi_2$*, is a representation of* $\mathfrak{g} \oplus \mathfrak{h}$ *acting on* $U \otimes V$*, given by* $(\pi_1 \otimes \pi_2)(X, Y) = \pi_1(X) \otimes I + I \otimes \pi_2(Y)$ *for all* $x \in \mathfrak{g}$ *and* $Y \in \mathfrak{h}$*.*

Note that $\pi_1 \otimes \pi_2$ in the above definition is indeed a representation.

We already defined a tensor product of a representation with a product group $G \times H$ acting on a space $U \otimes V$, but we can also define a tensor product of a representation with two different representations of the same group G acting on spaces U and V.

Definition 4.4.5. *Let G be a matrix Lie group and Π_1 and Π_2 be representations of G acting on spaces U and V. Then the* ***tensor product*** *representation of G acting on $U \otimes V$ is defined by $\Pi_1 \otimes \Pi_2)(A) = \Pi_1(A) \otimes \Pi_2(A)$ for all $A \in G$. Similarly, if π_2 and π_2 are representations of a Lie algebra $\mathfrak{g}$, we define a tensor product representation of $\mathfrak{g}$ on $U \otimes V$ by $(\pi_1 \otimes \pi_2)(X) = \pi_1(X) \otimes I + I \otimes \pi_2(X)$ for all $X \in \mathfrak{g}$.*

These are the essentials of tensor product representations, but we also have another interesting type of representation that we will discuss next: the dual representation.

4.5 Dual Representations

Much like with tensor products, we discussed dual spaces in the Background Material chapter so that we could jump right into dual space representations here.

Definition 4.5.1. *Let G be a matrix Lie group and $\Pi : G \to GL(V)$ be a representation. Then the* ***dual representation*** *is the representation $\Pi^* : G \to GL(V^*)$ defined by $\Pi^*(g)(f)(v) = f(\Pi(g^{-1})v)$. Similarly, if $\mathfrak{g}$ is a Lie algebra and $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation, then $\pi^* : \mathfrak{g} \to \mathfrak{gl}(V^*)$ is defined by $\pi^*(X)(f)(v) = f(\pi(-X)v)$.*

It can be checked that these are indeed representations. It is also a fact that if V is an inner product space, then for all $f \in V^*$, there exists $w \in V : f(v) = \langle w, v \rangle$, and so the functions are determined by the inner product. So you have a pairing $V \times V^*$ with $\langle v, f \rangle = f(v)$.

We also have the following result that we will not prove, but is very interesting nonetheless.

Theorem 4.5.2. *Let Π be a representation of a matrix Lie group G. Then Π^* is irreducible if and only if Π is irreducible. Also, $(\Pi^*)^* \cong \Pi$. There are analogous results for Lie algebra representations.*

4.6 Representations of $Lie(SL_2(\mathbb{C}))$

We will now begin to look at representations of $Lie(SL_2(\mathbb{C}))$ and (eventually) representations of $Lie(SL_3(\mathbb{C}))$. Recall that $Lie(SL_2(\mathbb{C}))$ is the set of 2×2 matrices with trace 0 and that it has dimension 3. So we choose the following as our basis:

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The commutation relations are,

$$[H, X] = 2X \qquad [H, Y] = -2Y \qquad [X, Y] = H.$$

Now if V is a finite-dimensional complex vector space and A, B, and C are operators on V satisfying the same commutation relations as H, X, and Y, then we have

$$[A, B] = 2B,$$

$$[A, C] = -2C,$$

$$[B, C] = A.$$

Due to the skew symmetry and bilinearity of the brackets, there is a unique linear map $\pi : Lie(SL_2(\mathbb{C})) \to Lie(GL(V))$.

We can now prove the following theorem.

Theorem 4.6.1. *For each integer $m \geq 0$, there is an irreducible representation of $Lie(SL_2(\mathbb{C}))$ with dimension $m+1$. Any two irreducible complex representations of $Lie(SL_2(\mathbb{C}))$ with the same dimension are isomorphic. If π is an irreducible complex representation of $Lie(SL_2(\mathbb{C}))$ with dimension $m + 1$, then π is isomorphic to the representation π_m.*

Proof. Let π be an irreducible representation of $Lie(SL_2(\mathbb{C}))$ acting on a finite dimensional complex vector space V. We want to diagonalize the operator $\pi(H)$. Since we're working over $\mathbb{C}$, $\pi(H)$ must have at least one eigenvector. This is because all roots to the auxiliary equation are in the complex field. Let u be an eigenvector with eigenvalue α. Then if we apply Theorem 4.6.3 repeatedly,

$$\pi(H)\pi(X)^k u = (\alpha + 2k)\pi(X)^k u.$$

Since we're working over a finite dimensional space, there are only a finitely many number of eigenvectors and eigenvalues. Since applying $\pi(X)$ takes us to a new vector with eigenvalue $\alpha + 2$ there has to come a point where we have reached the highest eigenvalue and any more repeated use of $\pi(X)$ cannot take you to another eigenvector, so we say it goes to 0. More explicitly, there exists some $N \geq 0$ such that,

$$\pi(X)^N u \neq 0$$

but,

$$\pi(X)^{N+1} u = 0.$$

Let's call this highest eigenvector $u_{high} = \pi(X)^N u$ and $\alpha_{high} = \alpha + 2N$. Then we have

$$\pi(H)u_{high} = \alpha_{high} u_{high},$$

$$\pi(X)u_{high} = 0.$$

Now since applying $\pi(X)$ raised the eigenvalue by 2 and we know that if we apply $\pi(Y)$ the eigenvalue will decrease by 2, this motivates us to define

$$u_k = \pi(Y)^k u_{high},$$

for $k \geq 0$. Now if we apply $\pi(H)$, which just tells us which eigenvalue we are at, we get

$$\pi(H)u_k = (\alpha_{high} - 2k)u_k.$$

Now if we apply $\pi(X)$ to u_k, then we get

$$\pi(X)u_k = k[\alpha_{high} - (k-1)]u_{k-1}.$$

Now by the same logic, since we're dealing with a finite-dimensional space, if we continuously apply $\pi(Y)$ (which lowers the eigenvalue), we will eventually cycle through all of them and any extra application will give 0. So there exists an integer m such that

$$u_k = \pi(Y)^k u_{high} \neq 0$$

for all $k \leq m$, but

$$u_{m+1} = \pi(Y)^{m+1} u_{high} = 0.$$

Now if $u_{m+1} = 0$, then $\pi(X)u_{m+1} = 0$. Therefore,

$$0 = \pi(X)u_{m+1} = (m+1)(\alpha_{high} - m) = 0.$$

Since u_m and $m+1$ are nonzero, this means $\alpha_{high} - m = 0$. So α_{high} coincides with the non-negative integer m. From this comes the conclusions that for every irreducible representation π, there exists an integer $m \geq 0$ and associated non-zero vectors u_{high}... u_0 that satisfy the following conditions:

1. $\pi(H)u_k = (\alpha_{high} - 2k)u_k$
2. $\pi(Y)u_k = \begin{Bmatrix} u_{k+1}, & \text{if } k < m \\ 0, & \text{if } k = m \end{Bmatrix}$
3. $\pi(X)u_k = \begin{Bmatrix} k(\alpha_{high} - (k-1))u_{k-1}, & \text{if } k > 0 \\ 0, & \text{if } k = m \end{Bmatrix}$

The eigenvectors must be linearly independent since they are eigenvalues of $\pi(H)$ with distinct eigenvalues. Also, the span of u_α ... u_0 is invariant under $\pi(H), \pi(Y)$, and $\pi(X)$. Therefore this is true for all $\pi(Z)$, for $Z \in Lie(SL_2(\mathbb{C}))$, since Z would be some combination of H, Y, and X. Since π is irreducible, the space must be all of V.

The preceding discussion has shown that every irreducible representation of $Lie(SL_2(\mathbb{C}))$ satisfies the previous conditions. Now if we defined $\pi(H)$, $\pi(Y)$, $\pi(X)$ by the previous conditions, then one finds that the operators satisfy the commutations relations we started with. Therefore any irreducible representation with dimension $m+1$ must satisfy the previous conditions, making them all isomorphic. □

Now were going to prove a nice corollary that comes from this theorem that shows all shows all eigenvalues for $\pi(H)$ are integers.

Corollary 4.6.2. *If π is a finite-dimensional representation of $Lie(SL_2(\mathbb{C}))$, not necessarily irreducible, then every eigenvalue of $\pi(H)$ is an integer and if v is an eigenvector for $\pi(H)$ with eigenvalue λ and $\pi(X)v = 0$, then λ is a non-negative integer.*

Proof. Suppose v is an eigenvector for $\pi(H)$ with an associated eigenvalue λ. There exists some $N \geq 0$ such that $\pi(X)^N v \neq 0$, but $\pi(X)^{N+1} v = 0$. Now $\pi(X)^N v$ is an eigenvector of $\pi(H)$ with eigenvalue $\lambda + 2N$ shown in the previous theorem, $m = \lambda + 2N$ must be a non-negative integer, which means λ must be an integer. If $\pi(X)v = 0$ then take $N = 0$ and $\lambda = m$ again is non-negative. □

4.7 Roots and Weights

In this section, we discuss the roots and weights of Lie algebras. Specifically we will define a weight and a root for $Lie(SL_3(\mathbb{C}))$, since that is the next example we are going to look at.

Definition 4.7.1. *If (π, V) is a representation of Lie($SL_3(\mathbb{C})$), then an ordered pair $\mu = (m_1, m_2) \in \mathbb{C}^2$ is called a weight for π if there exists $v \neq 0 \in V$ such that,*

$$\pi(H_1)v = m_1 v$$

$$\pi(H_2)v = m_2 v$$

We can now define a root, which is the weight of the adjoint representation.

Definition 4.7.2. *An ordered pair $\alpha = (\alpha_1, \alpha_2) \in \mathbb{C}^2$ is called a root if,*

1. *α_1, α_2 are not zero*

2. *there exists a nonzero $Z \in$ Lie($SL_3(\mathbb{C})$) such that*

$$[H_1, Z] = \alpha_1 Z$$

$$[H_2, Z] = \alpha_2 Z$$

Theorem 4.7.3. *Let $\alpha = (a_1, a_2)$ be a root and let $Z_\alpha \in$ Lie($SL_3(\mathbb{C})$) be a corresponding root vector. Let π be a representation of Lie($SL_3(\mathbb{C})$), $\mu = (m_1, m_2)$ be a weight for π, and let $v \neq 0$ be a corresponding weight vector. Then we have the following:*

$$\pi(H_1)\pi(Z_\alpha)v = (m_1 + a_1)\pi(Z_\alpha)v,$$

$$\pi(H_2)\pi(Z_\alpha)v = (m_2 + a_2)\pi(Z_\alpha)v.$$

Thus, either $\pi(Z_\alpha)v = 0$ or $\pi(Z_\alpha)v$ is a new weight vector with weight

$$\mu + \alpha = (m_1 + a_1, m_2 + a_2).$$

Proof. By the definition of a root, we have the commutation relation

$[H_1, Z_\alpha] = \pi(H_1)\pi(Z_\alpha)v - \pi(Z_\alpha)\pi(H_1) = a_1 Z_\alpha$. Thus,

$$\pi(H_1)\pi(Z_\alpha)v = (\pi(Z_\alpha)\pi(H_1) + a_1\pi(Z_\alpha))v$$

$$= \pi(Z_\alpha)(m_1 v) + a_1\pi(Z_\alpha)v$$

$$= (m_1 + a_1)\pi(Z_\alpha)v.$$

A similar argument allows us to compute $\pi(H_2)\pi(Z_\alpha)v$. □

Now we will define a notion of higher and lower weights.

Definition 4.7.4. *Let α_1 and α_2 be roots corresponding to the weights μ_1 and μ_2, then μ_1 is called **higher** than μ_2 if $\mu_1 - \mu_2$ can be rewritten as,*

$$\mu_1 - \mu_2 = a\alpha_1 + b\alpha_2$$

with $a \geq 0$ and $b \geq 0$.

Definition 4.7.5. *If π is a representation of $Lie(SL_3(\mathbb{C}))$, then a weight μ_0 is said to be a **highest weight** if for all weights μ, $\mu_0 \geq \mu$.*

Now, you may be wondering why we are interested in $Lie(SL_3(\mathbb{C}))$? Well its because $Lie(SL_3(\mathbb{C})) \simeq Lie(SU(3))$. This is because the special unitary matrices have a physical significance in physics. Particularly one can view $Lie(SU(n))$ as a set of rotation operators in n dimensions. This leads to the notion of spin that is of great importance in quantum mechanics.

4.8 Representations of $Lie(SL_3(\mathbb{C}))$

Looking at $Lie(SL_3(\mathbb{C}))$ is similar to when we looked at $Lie(SL_2(\mathbb{C}))$, but now we increase the intricacy of the problem by adding one more "H" matrix into the basis.

But the dimension of $Lie(SL_3(\mathbb{C}))$ is $3^2 - 1 = 8$ while $Lie(SL_2(\mathbb{C}))$ was $2^2 - 1 = 3$, so simply adding one more H is not enough; we also have to add two more "X" matrices and two more "Y" matrices. The basis we will choose for$Lie(SL_3(\mathbb{C}))$ is:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$Y_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad Y_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad Y_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

It should be noted that if we take a subalgebra with basis $\langle H_1, X_1, Y_1\rangle$ or $\langle H_2, X_2, Y_2\rangle$ you get a subalgebra isomorphic to $Lie(SL_2(\mathbb{C}))$. Therefore, the commutation relations for each subalgebra are the same as $Lie(SL_2(\mathbb{C}))$:

$$[H_1, X_1] = 2X_1, \qquad [H_2, X_2] = 2X_2$$

$$[H_1, Y_1] = -2Y_1, \qquad [H_2, Y_2] = -2Y_2$$

$$[X_1, Y_1] = H_1, \qquad [X_2, Y_2] = H_2$$

Now the rest of the commutation relations are

$$[H_1, H_2] = 0$$

$$[H_2, X_1] = -X_1, \qquad [H_2, Y_1] = Y_1$$

$$[H_1, X_2] = -X_2, \qquad [H_1, Y_2] = Y_2$$

$$[H_1, X_3] = X_3, \qquad [H_1, Y_3] = -Y_3$$

$$[H_2, X_3] = X_3, \qquad [H_2, Y_3] = -Y_3$$

$$[X_3, Y_3] = H_1 + H_2$$

$$[X_1, X_2] = X_3, \qquad [Y_1, Y_2] = -Y_3$$

$$[X_1, Y_2] = 0, \qquad [X_2, Y_1] = 0$$

$$[X_1, X_3] = 0, \qquad [Y_1, Y_3] = 0$$

$$[X_2, X_3] = 0, \qquad [Y_2, Y_3] = 0$$

$$[X_2, Y_3] = Y_1, \qquad [X_3, Y_2] = X_1$$

$$[X_1, Y_3] = -Y_2, \qquad [X_3, Y_1] = -X_2$$

Since the commutator between H_1 and H_2 is 0, this means they share an eigenbasis and are able to be simultaneously diagonalized. Similar to $Lie(SL_2(\mathbb{C}))$, we will look for a representation that satisfies the same commutation relations. So we want to simultaneously diagonalize $\pi(H_1)$ and $\pi(H_2)$. Now because we have two H matrices, this means we will have two weights. However we know nothing about $\pi(H_i)v$, so we will look at the roots for this space.

Table 4.1: Roots for $Lie(SL_3(\mathbb{C}))$

α	Z	α	Z
(2,-1)	X_1	(-2,1)	Y_1
(-1,2)	X_2	(1,-2)	Y_2
(1,1)	X_3	(-1,-1)	Y_3

Identifying what the representations of $Lie(SL_3(\mathbb{C}))$ look like will be done in the proof of a theorem that has five different parts. But first, we look at a definition!

Definition 4.8.1. *A representation π of $Lie(SL_3(\mathbb{C}))$ is said to be a **highest weight cyclic representation** with weight $\mu = (m_1, m_2)$ if there exists $v \in V$ and $v \neq 0$ such that:*

1. *v is a weight vector with weight μ*
2. *$\pi(X_j)v = 0$ for $j = 1, \quad 2, \quad 3$*
3. *The smallest invariant subspace containing v is V*

Now we can look at the five-part theorem we mentioned.

Theorem 4.8.2. *We have the following:*

1. *Every irreducible representation of $Lie(SL_3(\mathbb{C}))$ is the direct sum of its weight spaces.*
2. *Every irreducible representation of $Lie(SL_3(\mathbb{C}))$ has a unique highest weight.*
3. *Two irreducible representations of $Lie(SL_3(\mathbb{C}))$ with the same highest weight are isomorphic.*
4. *The highest weight μ of an irreducible representation must be of the form $\mu = (m_1, m_2)$, where m_1 and m_2 are non-negative integers.*
5. *For every pair (m_1, m_2) of non-negative integers, there exists an irreducible representation of $Lie(SL_3(\mathbb{C}))$ with the highest weight (m_1, m_2).*

Proof. Since the theorem is divided into 5 parts, we will divide our proof into 5 parts.

1. Let W be the sum of weight spaces in V. So $W = \{w \in V : w = a_1v_1 + a_2v_2 + ... + a_nv_n$ where $\pi(H_1)v_i = \lambda_i v_i$ and $\pi(H_2)v_i = \lambda_i v_i\}$. Now every representation of $Lie(SL_3(\mathbb{C}))$ has at least one weight. From this we know $W \neq 0$. Now, from earlier we defined Z_α which took the weight space corresponding to μ and rotated it in the $\mu + \alpha$ weight space. Since W is made

up of all the weight spaces of V, any Z_α acting on W will be invariant since it will give back another weight space which is inside W. Now remember that $X_1, \quad X_2, \quad X_3, \quad Y_1, \quad Y_2, \quad Y_3$ are Z_α terms. Also, because H_1 and H_2 just tell you the eigenspace you're already in, W is invariant on those terms as well. So W is invariant under the entire basis of $Lie(SL_3(\mathbb{C}))$ and, therefore, $W = V$.

2. We have just shown that every irreducible representation of $Lie(SL_3(\mathbb{C}))$ is a direct sum of its weight spaces. Remember that the representation has finite dimension, this means that there can only be a finite number of weights. This means there must be a highest weight μ. A highest weight means for any weight vector v then,

$$\pi(X_j)v = 0 \qquad j = 1, \quad 2, \quad 3$$

Now since π is irreducible, that means the smallest invariant subspace that contains v must be the entire space.

3. Suppose π and σ are both irreducible representations with the same highest weight μ, V is the vector space corresponding to the representation π, W is the vector space corresponding to σ, and let v and w be the highest weight vectors from V and W, respectively. Now, to actually prove this, there is something that needs to be known about completely reducible representations of $Lie(SL_3(\mathbb{C}))$. That is if there is a completely reducible representation of $Lie(SL_3(\mathbb{C}))$ that is also highest weight cyclic, then it is irreducible. Now consider a representation $V \oplus W$ and let U be the smallest invariant subspace of $V \oplus W$ which contains (v, w). By definition, U is a highest weight cyclic representation. Now, because $V \oplus W$ is completely reducible, then U is also completely reducible. This means U is also irreducible. Now consider two maps, called projection maps P_v and

P_w. By projection, we mean that given a vector (v, w):

$$P_v(v, w) = v \qquad\qquad P_w(v, w) = w.$$

Now P_v and P_w are intertwining maps. This is also true if we restrict them to U instead of $V \oplus W$. Since U is the smallest invariant subspace that contains (v, w), then $P_v(v, w) = v$ and $P_w(v, w) = w$ is not the 0 map. Therefore, by Schur's Lemma [4], P_v is an isomorphism from U to V and P_w is an isomorphism from U to W, which means $V \simeq U \simeq W$. Therefore, V and W are isomorphic, where the isomorphism between V and W is $P_v(P_w^{-1})$.

4. Now if we restrict π to $\{X_1, Y_1, H_1\}$ or $\{X_2, Y_2, H_2\}$, then it is isomorphic to $Lie(SL_2(\mathbb{C}))$. Therefore, with these restrictions we know that m_1 and m_2 must be integers. similarly, we know that they are non-negative in $Lie(SL_2(\mathbb{C}))$, therefore they must be non-negative here as well. For if they were non-integer or negative, then when we applied the restrictions, which are isomorphic to $Lie(SL_2(\mathbb{C}))$, there would be a contradiction.

5. Let V_1 be the standard representation of $Lie(SL_3(\mathbb{C}))$, which will have weight vectors e_1, e_2, and e_3, which gives the weights $(1, 0)$, $(-1, 1)$, and $(0, -1)$, respectively. The dual of the standard representation is $\pi(Z) = -Z^T$ for all $Z \in Lie(SL_3(\mathbb{C}))$, and has the weights $(-1, 0)$, $(1, -1)$, $(0, 1)$. The highest weight is $(0, 1)$. We call this space V_2. Now were going to build a general weight space of highest weight (m_1, m_2) with various combinations of V_1 and V_2. Now consider π_{m_1,m_2} which is given by

$$(V_1 \otimes V_1 \otimes ... \otimes V_1) \otimes (V_2 \otimes V_2 \otimes ... \otimes V_2)$$

where there are m_1 V_1's and m_2 V_2's. Then $v = v_1 \otimes ... \otimes v_1 \otimes v_2 \otimes ... \otimes v_2$ is a vector with weight (m_1, m_2). Let W be the smallest invariant subspace containing v. If π_{m_1,m_2} is completely reducible, then W will be completely reducible. Then W is irreducible since it is highest weight cyclic. This means W is an irreducible space that has highest weight (m_1, m_2).

□

We have now finished our discussion on $Lie(SL_3(\mathbb{C}))$. Before diving into some physical applications of this work, we will take a brief detour into discussing the Weyl Group, as this is also an interesting aspect of Lie theory.

4.9 The Weyl Group

In this section, we talk about the Weyl group. There are many interesting directions one could take with this subject, although we will only discuss some aspects of this engaging topic. We begin by letting $\mathfrak{h} = span\{H_1, H_2\}$, where

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Note that $\mathfrak{h}$ is a two-dimensional subspace of $Le(SL_3(\mathbb{C}))$. Now, for $A \in SU(3)$ and $H \in h$, we have $Ad_A(H) = AHA^{-1}$. We can then define two more groups: let $N = \{A \in SU(3) : Ad_A(H) \in h \text{ for all } H \in h\}$ and $Z = \{A \in SU(3) : Ad_A(H) = H \text{ for all } H \in h\}$. Note that $AHA^{-1} = H \implies AH = HA$, and so we can think of Z as a kind of center (as in, the center of a group), hence its label of Z. We then have a few results that follow from these definitions, all of which involve basic group theory.

Theorem 4.9.1. *N is a subgroup of $SU(3)$.*

Proof. By definition, N is a subset of $SU(3)$. So we just need to show that N is a group (under the same operation as $SU(3)$, which is multiplication). Let x,$y \in N$ and $H \in h$.

1. Closure: We have $Ad_{xy}(H) = (xy)H(xy)^{-1} = xyHy^{-1}x^{-1} = xAd_y(H)x^{-1}$. Note that since $y \in N$, $Ad_y(H) \in h$. Thus, $xAd_y(H)x^{-1} = Ad_x(Ad_y(H))$, which is also in h because $x \in N$. Thus, $xy \in N$ and we have closure.

2. Associativity: This follows from the associativity of $SU(3)$.

3. Identity: Note that the 3×3 identity matrix is in $SU(3)$, as it is unitary and has determinant 1. Thus, if we let e be said identity matrix, we have $Ad_e(H) = eHe^{-1} = H \in h$. So $e \in N$ and we have an identity element.

4. Inverse:We have $Ad_x(H) = H' \in h \implies xHx^{-1} = H' \implies H = x^{-1}H'x \implies H = Ad_{x^{-1}}(H')$. Since H' is arbitrary, $x^{-1} \in N$ and we have inverses.

   Thus, N is a subgroup of $SU(3)$.

□

Theorem 4.9.2. *Z is a subgroup of $SU(3)$.*

Proof. By definition, Z is a subset of $SU(3)$ (and actually N, as $H \in h$). So we just need to show that Z is a group (under the same operation as $SU(3)$, which is multiplication). Let x,$y \in Z$ and $H \in h$.

1. Closure: We have $Ad_{xy}(H) = (xy)H(xy)^{-1} = xyHy^{-1}x^{-1} = xAd_y(H)x^{-1}$. Note that since $y \in Z$, $Ad_y(H) = H$. Thus, $xAd_y(H)x^{-1} = xHx^{-1} = Ad_x(H) = H$. Thus, $xy \in Z$ and we have closure.

2. Associativity: This follows from the associativity of $SU(3)$.

3. Identity: Note that the 3×3 identity matrix is in $SU(3)$, as it is unitary and has determinant 1. Thus, if we let e be said identity matrix, we have $Ad_e(H) = eHe^{-1} = H$. So $e \in Z$ and we have an identity element.

4. Inverse: We have $Ad_x(H) = H \implies xHx^{-1} = H \implies H = x^{-1}Hx \implies H = Ad_{x^{-1}}(H)$. Thus, $x^{-1} \in Z$ and we have inverses.

   Thus, Z is a subgroup of $SU(3)$.

□

Theorem 4.9.3. *Z is a normal subgroup of N.*

Proof. As mentioned in the last proof, Z is a subset of N and a subgroup of $SU(3)$. Since N and $SU(3)$ have the same operation, it follows that Z is a subgroup of N. To show normality, we want to show that $nzn^{-1} \in Z$ for all $z \in Z$ and $n \in N$. So let $z \in Z$, $n \in N$, and $H \in h$. Then we have $Ad_{nzn^{-1}}(H) = nzn^{-1}Hnz^{-1}n^{-1}$. Since $n \in N$, this equals $nzH'z^{-1}n^{-1}$, where $H' \in h$. Since $z \in Z$, this then equals $nH'n^{-1}$, which must equal H (since we had $n^{-1}Hn = H'$). Thus, $nzn^{-1} \in Z$ and Z is a normal subgroup of N. □

Since Z is a normal subgroup of N, we can make a quotient group that will be denoted by $W = \frac{N}{Z}$. This is known as the **Weyl group**. Now, W acts on h in the following way: for $w = [A] \in W$, where $A \in N$, we have $w.H = Ad_A(H)$. This action is well-defined since if B is an element of the same coset as A, then $B = AC$ (where $C \in Z$) and we get $Ad_B(H) = Ad_A(Ad_C(H)) = Ad_A(H)$ since $C \in Z$. We now want to prove the following theorem:

Theorem 4.9.4. *This theorem states that we are able to redefine Z and N as*

$$Z = \{\begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{i\phi} & 0 \\ 0 & 0 & e^{-i(\theta+\phi)} \end{pmatrix} : \theta, \phi \in \mathbb{R}\},$$

which we will denote as Z'*, and* $N = \{A \in SU(3)$: *for all* $j \in \{1,2,3\}, \exists k_j \in \{1,2,3\}, \theta_j \in \mathbb{R} : Ae_j = e^{i\theta_j} e_{k_j}\}$. *(Note that* e_1, e_2, e_3 *is the standard basis for* $\mathbb{C}^3$.*) This then implies that* $W = \frac{N}{Z} \simeq S_3$.

Before we go through the proof of this, it is a good idea to list an example of such a matrix in N, as the new definition looks very scary. Such a matrix would look like

$$\begin{pmatrix} 0 & 0 & e^{i\theta_3} \\ e^{i\theta_1} & 0 & 0 \\ 0 & e^{i\theta_2} & 0 \end{pmatrix},$$

where $\theta_1 + \theta_2 + \theta_3 = 0$. This is so the matrix is invertible and has determinant 1, which is necessary since it is also an element of $SU(3)$. Now we continue with the proof of the theorem.

Proof. The first task that we want to conquer is to find Z. Note that if we pick an element in Z', this is also an element of $SU(3)$, as it is unitary with determinant 1, and is satisfies the adjoint condition. So we know that $Z' \subseteq Z$. Now, suppose that $A \in Z$. Then since $H_1 \in h$, we have $AH_1 = H_1 A$. Now, $H_1 e_1 = 1e_1$. So $H_1 A e_1 = A H_1 e_1 = A 1 e_1 = 1 A e_1$. This means that Ae_1 is an eigenvector for H_1 with an eigenvalue of 1. So $Ae_1 \in E_1 = span\{e_1\}$. So there exists $\lambda \in \mathbb{C} : Ae_1 = \lambda e_1$. This works similarly for e_2 and e_3. So A is diagonal and thus we have $A \in Z'$. Thus, $Z \subseteq Z'$. Thus, we have found Z.

Now we want to find N. Suppose that $A \in N$. Then $AH_1A^{-1} \in h$. So AH_1A^{-1} is diagonal. So e_1, e_2, e_3 are eigenvectors for AH_1A^{-1}. Note that $AH_1A^{-1}(Ae_1) = AH_1e_1 = A1e_1 = 1Ae_1$. So Ae_1 is an eigenvector for AH_1A^{-1}. So there exists $j : A \in span\{e_j\}$. Note that if $Av = \lambda v$ and $A \in U(n)$, then $AA^* = A^*A = I$ and so $v^*A^* = v^*\lambda^* = \lambda^* v^*$ (as λ is a scalar, so it commutes). This then implies that $\|v\|^2 = v^*v = v^*Iv = v^*A^*Av = \lambda^* v^* Av = \lambda^* v^* \lambda v = \lambda^* \lambda v^* v = |\lambda|^2 \|v\|^2$. This implies that $|\lambda|^2 = 1$, which implies that $|\lambda| = 1$. Thus, there exists $\theta \in \mathbb{R} : \lambda = e^{i\theta}$.

Finally, we want to show that $W = \frac{N}{Z} \simeq S_3$. Note that for all $A \in N$, we get a bijective linear transformation (since A is an invertible matrix), which we can denote by $T_A : \mathbb{C}^3 \to \mathbb{C}^3$ and which is defined by $T_A(v) = Av$. So $T_A(e_i) = \lambda e_j$, where $\lambda \in \mathbb{C}$. We then get a surjective homomorphism that maps from N to $S_{\{e_1,e_2,e_3\}} = S_3$ defined by $A \to T_A$. The kernel of this homomorphism is $\{A : T_A = id\}$, which is Z. Thus, by the First Isomorphism Theorem, we have that $\frac{N}{Z} \simeq S_3$. □

Now we can define an inner product on h by $\langle H, H' \rangle = tr(H^* H')$. Note that this is the Hilbert-Schmidt inner product given in the Background Material chapter. We will use this in our next definition.

Definition 4.9.5. *For a representation π of $Lie(SL_3(\mathbb{C}))$, $\lambda \in h$ is a **weight** for π if there exists a nonzero $v \in V$ such that $\pi(H)v = \langle \lambda, H \rangle v$ for all $H \in h$. We call v the **weight vector**, as it is an eigenvector.*

We now have the following result:

Theorem 4.9.6. *Let Π be a representation of $SU(3)$ and π be a representation of $Lie(SL_3(\mathbb{C}))$. If λ is a weight for π, then $w.\lambda$ is also a weight for π for all $w \in W$. Also, λ and $w.\lambda$ have the same multiplicity, where the multiplicity of a weight is equal to the dimension of a weight space.*

Proof. First note that for $H, H' \in h$, $w = [u] \in \frac{N}{Z}(u \in N) = uz$, a coset. Then we have $\langle w.H, H' \rangle = \langle uHu^{-1}, H' \rangle = tr((uHu^{-1})^* H') = tr(uHu^{-1}H')$, where we are using the fact that we have diagonal and unitary matrices. We also have that $\langle H, w^{-1}.H' \rangle = tr(H^* u^{-1} H' u) = tr(uH^{-1}u^{-1}H') = \langle w.H, H' \rangle$.

Now, say λ is a weight with weight vector v. Then for al $u \in N, H \in h$, we have $\pi(H)\Pi(u)v = \Pi(u)\Pi(u^{-1})\pi(H)\Pi(u)v = \Pi(u)\pi(u^{-1}Hu)v = \Pi(u)\langle \lambda, u^{-1}Hu \rangle v$, where we note that $u^{-1}Hu \in h$. This then equals $\Pi(u)\langle \lambda, w^{-1}.H \rangle v$, where $w = [u]$, which then equals $\langle w.\lambda, H \rangle \Pi(u)v$. So $w.\lambda$ is a weight with weight vector $\Pi(u)v$. Now, recall that $\Pi(u)$ is an invertible linear map from V to itself. So we also have

$\Pi(u) : E_\lambda \to E_{w.\lambda}$, which can also have an inverse map $\Pi(u^{-1})$. So we have $E_\lambda \simeq E_{w.\lambda}$, and so $dim(E_\lambda) = dim(E_{w.\lambda})$. $\square$

Chapter 5

Physics Applications

This chapter covers some brief physical applications of the mathematical work we have looked at, as well as context necessary for those without a physics background. We will go through an example of constructing an irreducible representation of highest weight $\frac{1}{2}$ using a more physics-based approach (although it will still be mathematically rigorous) and explain what that means from a physics perspective, although we will need to go over some facts and define certain objects first.

5.1 The Rotation Group

A very important note is that $SO(3)$ is known as the **rotation group** because it is the group that models the rotations in a three-dimensional space; that is, if you take the origin of $\mathbb{R}^3$, then $SO(3)$ is the group of the rotations about that origin. So $SO(3)$ is an important group because it is one that represents the physical world in which we live. We ignore the details of why this is the case, but it is common knowledge for physicists, so we accept it.

Now, we know that $SO(3)$ is a Lie group and $Lie(SO(3))$ is its corresponding Lie algebra. We also know that $Lie(SO(3)) \cong Lie(SU(2))$. What we are going to do in the next section is construct an irreducible representation of $Lie(SU(2))$ with a highest weight of $\frac{1}{2}$. While we will not go in-depth with the details, note that this is useful in physics because of the connection we just made: $Lie(SU(2)) \cong Lie(SO(3))$, which is the corresponding Lie algebra of $SO(3)$, the rotation group that models the real world we live in.

5.2 The Physics Approach

We call this section the Physics Approach not because we are necessarily using a lot of physics information - it is still mainly pure mathematics - but because this specific way of going about it is how a physicist would. The notation would vary slightly and some of the details might not be figured out with as much mathematical precision, but the basic form is very much what a physicist would see. So let's begin!

Note the following information, some of which we have mentioned before, all of which comes from Hall's book [4]:

- For compact Lie groups, their representations are isomorphic to unitary representations.

- If a Lie group representation Π is unitary, then its corresponding Lie algebra representation π satisfies $\pi(g)^* = -\pi(g)$ for all g in the Lie algebra (note that we discussed this earlier).

- We know that $SU(2)$ is compact (and we can ignore what exactly being compact means). So $\Pi : SU(2) \mapsto GL_n(\mathbb{C})$ is always unitary or isomorphic to a unitary representation, and so the Lie algebra representation corresponding to Π, $\pi : Lie(SU(2) \mapsto M_n(\mathbb{C})$, either satisfies $\pi(g)^* = -\pi(g)$ for all $g \in Lie(SU(2))$, or π is isomorphic to a representation with this skew self-adjoint property.

- We also know that if a Lie group G is simply connected, then all representations π of $Lie(G)$ come from the representations Π of G.

- Since we know that $SU(2)$ is simply connected, all $Lie(SU(2))$ representations π come from $SU(2)$ representations. Thus, for a Lie algebra representation $\pi : Lie(SU(2)) \mapsto M_n(\mathbb{C})$, we always have $\pi(g)^* = -\pi(g)$ for all $g \in Lie(SU(2))$.

Let $\pi : Lie(SU(2)) \mapsto M_n(\mathbb{C})$ be a Lie algebra representation. We choose our basis for $Lie(SU(2))$ to be the following matrices:

$$\sigma_1 = \begin{pmatrix} \frac{i}{2} & 0 \\ 0 & \frac{-i}{2} \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & \frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 0 & \frac{-1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.$$

Note that this is the same basis we chose earlier when we worked on this, way back when, and so we know that these satisfy the commutation relations: $[\sigma_1, \sigma_2] = \sigma_3, [\sigma_2, \sigma_3] = \sigma_1$, and $[\sigma_3, \sigma_1] = \sigma_2$.

By our bullet notes above, we have that $\pi(\sigma_k)^* = -\pi(\sigma_k)$. Since π is a representation, we then get $-\pi(\sigma_k) = \pi(-\sigma_k)$. Finally, since $\sigma_k \in Lie(SU(2))$, we end up with $\pi(-\sigma_k) = \pi(\sigma_k^*)$. So $\pi(\sigma_k)^* = \pi(\sigma_k^*)$ Now define $J_k = -i\pi(\sigma_k)$, where $k \in \{1, 2, 3\}$. Then we get that $J_k^* = (-i\pi(\sigma_k))^* = i\pi(\sigma_k)^* = i\pi(\sigma_k^*) = i\pi(-\sigma_k) = -i\pi(\sigma_k) = J_k$, which means that J_k is Hermitian.

Now let's go over some commutation relations. We have $[J_1, J_2] = J_1J_2 - J_2J_1 = \pi(\sigma_1)\pi(\sigma_2) - \pi(\sigma_2)\pi(\sigma_1) = [\pi(\sigma_1), \pi(\sigma_2)] = \pi([\sigma_1, \sigma_2]) = \pi(\sigma_3) = iJ_3$. Repeating this process, we find that $[J_2, J_3] = iJ_1$ and $[J_3, J_1] = iJ_2$.

We can also define $J^2 = J_1^2 + J_2^2 + J_3^2$. Then $[J^2, J_k] = 0$. For example,

$$\begin{aligned}
[J^2, J_1] &= J^2J_1 - J_1J^2 \\
&= (J_1^2 + J_2^2 + J_3^2)J_1 - J_1(J_1^2 + J_2^2 + J_3^2) \\
&= J_1^3 + J_2^2J_1 + J_3^2J_1 - J_1^3 - J_1J_2^2 - J_1J_3^2 \\
&= J_2^2J_1 - J_1J_2^2 + J_3^2J_1 - J_1J_3^2 \\
&= J_2^2J_1 - J_2J_1J_2 + J_2J_1J_2 - J_1J_2^2 + J_3^2J_1 - J_3J_1J_3 + J_3J_1J_3 - J_1J_3^2 \\
&= J_2[J_2, J_1] + [J_2, J_1]J_2 + J_3[J_3, J_1] + [J_3, J_1]J_3
\end{aligned}$$

$$= -iJ_2J_3 - iJ_3J_2 + iJ_3J_2 + iJ_2J_3$$
$$= 0.$$

Also note that $[J^2, J_k] = 0 \implies J^2 J_k - J_k J^2 = 0 \implies J^2 J_k = J_k J^2$.

Now we will define $J_\pm = J_1 \pm iJ_2$. Then,

$$J_+J_- = (J_1 + iJ_2)(J_1 - iJ_2) = J_1^2 - iJ_1J_2 + iJ_2J_1 + J_2^2.$$

Similarly, $J_-J_+ = J_1^2 + iJ_1J_2 - iJ_2J_1 + J_2^2$. So adding these together gives us

$$J_+J_- + J_-J_+ = 2J_1^2 + 2J_2^2.$$

Thus, we have

$$J^2 = J_1^2 + J_2^2 + J_3^2 = \tfrac{1}{2}(J_+J_- + J_-J_+) + J_3^2.$$

Also,

$$[J_+, J_-] = J_+J_- - J_-J_+ = J_1^2 - iJ_1J_2 + iJ_2J_1 + J_2^2 - (J_1^2 + iJ_1J_2 - iJ_2J_1 + J_2^2) =$$
$$-2iJ_1J_2 + 2iJ_2J_1 = 2i[J_2, J_1] = 2J_3.$$

Next, we have

$$[J_+, J_3] = (J_1 + iJ_2)J_3 - J_3(J_1 + iJ_2) = J_1J_3 + iJ_2J_3 - J_3J_1 - iJ_3J_2 =$$
$$[J_1, J_3] + i[J_2, J_3] = -iJ_2 - J_1 = -(J_1 + iJ_2) = -J_+.$$

Similarly, $[J_-, J_3] = J_-$. Thus, we have that $[J_\pm, J_3] = \mp J_\pm$.

Moving forward, recall that we said $[J^2, J_k] = 0$ for $k \in \{1, 2, 3\}$. So $[J^2, J_3] = 0$, or $J^2 J_3 = J_3 J^2$. If we presume that we have distinct eigenvalues, we have

$$J^2 v = \lambda v$$
$$J_3 v = mv.$$

This follows from a theorem we showed in Chapter 2, although we note that even if we presumed eigenvalues were not distinct, this would still work. However, we omit the details of this.

Note that $[J_\pm, J_3] = \mp J_\pm \implies J_\pm J_3 - J_3 J_\pm = \mp J_\pm \implies J_3 J_\pm = \pm J_\pm + J_\pm J_3$. So $J_3 J_\pm v = (\pm J_\pm + J_\pm J_3)v = \pm J_\pm v + J_\pm J_3 v = \pm J_\pm v + J_\pm m v = (\pm 1 + m) J_\pm v = (m \pm 1) J_\pm v$. Therefore, we have that $J_3 J_\pm v = (m \pm 1) J_\pm v$.

We also have $J^2 J_\pm v = J^2 (J_1 \pm i J_2) v = (J^2 J_1 \pm i J^2 J_2) v = (J_1 J^2 \pm i J_2 J^2) v = (J_1 \pm i J_2) J^2 v = J_\pm \lambda v = \lambda J_\pm v$. Thus, we have $J^2(J_\pm v) = \lambda (J_\pm v)$.

So we have

$$J^2(J_\pm v) = \lambda(J_\pm v)$$
$$J_3 J_\pm v = (m \pm 1) J_\pm v.$$

Recall that any nonzero vector can be normalized. So

$$\langle J_k v, J_k v \rangle \geq 0 \implies (J_k v)^* J_k v \geq 0 \implies v^* J_k^* J_k v \geq 0 \implies v^* J_k^2 v \geq 0$$

since we know that J_k is Hermitian (and therefore $J_k = J_k^*$). Utilizing this, we get that

$$v^* J^2 v = v^*(J_1^2 + J_2^2 + J_3^2) v \implies v^* \lambda v = v^* J_1^2 v + v^* J_2^2 v + v^* J_3^2 v$$
$$\implies \lambda v^* v = v^* J_1^2 v + v^* J_2^2 v + m^2 v^* v$$

since $v^* J_3^2 v = v^* J_3 J_3 v = v^* J_3 m v = m v^* J_3 v = m v^* m v = m^2 v^* v$. We know that v can be normalized (since v is an eigenvector and therefore nonzero), so we have that $v^* v = \langle v, v \rangle = 1$. This means that we end up with $\lambda = v^* J_1^2 v + v^* J_2^2 v + m^2$. Now, $v^* J_1^2 v \geq 0$ and $v^* J_2^2 v \geq 0$ because we said that $v^* J_k^2 v \geq 0$. So we finally end up with $\lambda = v^* J_1^2 v + v^* J_2^2 v + m^2 \geq m^2 \implies \lambda \geq m^2 \implies$
$\lambda - m^2 \geq 0$.

Note that since $Lie(SU(2))$ is a finite-dimensional vector space, there are a finite number of eigenvalues. So if we start with $J_3 J_+ w = (m+1) w$ (where w is our eigenvector), we can apply J_+ on w a finite number (say n) times, eventually giving us $J_3(J_+)^n w = (m+n)(J_+)^n w$. Since n was our maximum finite number, we could not have $J_3(J_+)^{n+1} w = (m+n+1)(J_+)^{n+1} w$, where $(J_+)^{n+1} w$ is an eigenvector. But we know this formula still needs to be true, which means that $(J_+)^{n+1} w = 0$ because that eliminates it from being an eigenvector.

With this all being said, we can let $j = m+n$ such that $J_3(J_+)^n w = j(J_+)^n w$. Using our earlier terminology, j would be a **highest weight**. If let $v = (J_+)^n w$ for the sake of easier notation, we have that $J_+ v = 0$. This implies that $J_- J_+ v = 0$. So we get $0 = J_- J_+ v = (J_1 - iJ_2)(J_1 + iJ_2)v = (J_1^2 + iJ_1 J_2 - iJ_2 J_1 + J_2^2)v = (J^2 - J_3^3 + i[J_1, J_2])v = (J^2 - J_3^2 - J_3)v = J^2 v - J_3^2 v - J_3 v = \lambda v - j^2 v - jv = (\lambda - j^2 - j)v$. So $0 = (\lambda - j^2 - j)v$. Since v is an eigenvector, $v \neq 0$, and so we must have that $\lambda - j^2 - j = 0$, which implies that $\lambda = j^2 + j = j(j+1)$.

We have a similar result by letting $j' = m - n$ such that $J_3(J_-)^n w = j'(J_-)^n w$. Then we say j' is a **lowest weight**. Similar to what we did with j, we get that $\lambda = j'(j'-1)$. Thus, we have that $j(j+1) = j'(j'-1) \implies j = -j'$ or $j = j'-1$. But j' is the smallest weight, so $j \neq j'-1$, as that implies that $j'-1$ is the smallest. So we get that $j = -j'$. Since J_- lowers the value of m by 1, it is an integer number of times to go from j to $-j$. Since we are constantly subtracting by 1, this gives us a total of $2j$ steps (e.g., to go from 5 to -5, you have to subtract 1 ten times, or $2 * 5$). But we said this was an integer number, and so j must be a half-integer multiple. Thus, our highest and lowest weights must be half-integer multiples.

Now we note that $J_+^* = (J_1 + iJ_2)^* = J_1^* - iJ_2^* = J_1 - iJ_2$ (since we know that J_k is Hermitian). But this is just J_-. Similarly, we have that $J_-^* = J_+$. Thus, we have $J_\pm^* = J_\mp$.

We also know that since $J_\pm v$ is a nonzero vector, we can normalize it to get $J_\pm v = c_\pm w$, where $c_\pm$ is a scalar and w is a normalized vector.

We now have $v^* J_- J_+ v = (J_+ v)^* J_+ v = (c_+ w)^* c_+ w = w^* c_+^* c_+ w = c_+^* c_+ w^* w =$

$c_+^* c_+$ (since w is normalized). But this just equals $|c_+|^2$ since $|a| = \sqrt{a^* a}$ for a vector a.

But we also have that $v^* J_- J_+ v = v^*(J^2 - J_3^2 - J_3)v = v^*(J^2 v - J_3^2 v - J_3 v) = v^* \lambda v - v^* m^2 v - v^* m v = \lambda - m^2 - m$ (since we could choose v to be normalized, and therefore we get $v^* v = 1$ once we move around the scalars). But $\lambda = j(j+1)$, so we get $j(j+1) - m^2 - m$. So we end up with $|c_+|^2 = j(j+1) - m^2 - m \implies |c_+| = \pm\sqrt{j(j+1) - m^2 - m}$
$\implies |c_+| = \sqrt{j(j+1) - m^2 - m}$ since $|c_+|$ must be positive. Then we get that $c_+ = \pm\sqrt{j(j+1) - m^2 - m}$. But $|c_+| = \sqrt{c_+^* c_+}$, and so our square root must be positive. So we finally, we end up with $c_+ = \sqrt{(j-m)(j+m+1)}$. Similarly, $c_- = \sqrt{(j+m)(j-m+1)}$. Thus, we have that

$$J_\pm v = c_\pm w \implies J_\pm v = \sqrt{(j \mp m)(j \pm m + 1)}w.$$

We have shown a lot of mathematics, and so now we will show an example utilizing what we have learned.

5.3 Example with Highest Weight of $\frac{1}{2}$

Say we have a highest weight of $j = \frac{1}{2}$ (which we are allowed since we said they must be half-integer multiples). Then the lowest weight is $-j = -\frac{1}{2}$. Since m values go up and down by 1 each time, and we know it cannot get any higher than $\frac{1}{2}$ (or lower than $-\frac{1}{2}$), we know that the only possible values for m are $\frac{1}{2}$ and $-\frac{1}{2}$. Then we have that, for a vector v,

$$J_3 v = \tfrac{1}{2} v$$
$$J_3 w = -\tfrac{1}{2} w.$$

Now, let

$$v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } w = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

since they are easy enough matrices to work with, and try to construct a representation from here. Then since v and w are 2×1 matrices, J_3 must be a 2×2 matrix (for matrix multiplication to work). We have that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \implies \begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \implies a = \frac{1}{2} \text{ and } c = 0.$$

Similarly, we have that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{2} \end{pmatrix} \implies b = 0 \text{ and } d = -\frac{1}{2}.$$

Thus,

$$J_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Since $\frac{1}{2}$ is the highest weight, we know that $J_+v = 0$. Also, $J_3 w = -\frac{1}{2}w$ and we know that $J_3 J_+ w = (m+1)J_+ w$. In this case, $m = -\frac{1}{2}$, so $J_3 J_+ w = \frac{1}{2} J_+ w$. Since we presumed eigenvalues were distinct, it must be the case that $J_+ w = c_+ v$, where c_+ is a scalar that has a formula we know. So we get $J_+ w = \sqrt{(j-m)(j+m+1)}v = \sqrt{(\frac{1}{2} - (-\frac{1}{2}))(\frac{1}{2} + (-\frac{1}{2}) + 1)}v = v$.

So far, we have

$$J_+ v = 0$$
$$J_+ w = v.$$

This means that

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies e = g = 0$$

and

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \implies f = 1 \text{ and } h = 0.$$

Thus, we have

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

So

$$J_- = J_+^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Since $J_+ = J_1 + iJ_2$ and $J_- = J_1 - iJ_2$, we have:

$$J_1 = J_+ - iJ_2 \implies J_1 - iJ_2 = J_+ + iJ_2 - iJ_2 \implies J_- = J_+ - 2iJ_2 \implies J_- - J_+ = -2iJ_2.$$

We then get:

$$J_2 = \frac{J_+ - J_-}{2i} = \frac{1}{2i} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & \frac{1}{i} \\ \frac{-1}{i} & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ since } \frac{1}{i} = -i.$$

Finally, we end up with

$$J_1 = J_+ - i\frac{J_+ - J_-}{2i} = \frac{2iJ_+ - iJ_+ + iJ_-}{2i} = \frac{J_+ + J_-}{2} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Thus, we have found J_1, J_2, and J_3 given a highest weight of $\frac{1}{2}$. We can go even further and find out how the representation works, which is (from a mathematics perspective) what we would really be interested in.

We have the following:

$$-i \begin{pmatrix} 0 & \frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = J_1 = -i\pi(\sigma_1) \implies \pi(\sigma_1) = \begin{pmatrix} 0 & \frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix} \implies \pi(\sigma_1) = \sigma_2.$$

$$-i\begin{pmatrix}0 & \frac{1}{2}\\ -\frac{1}{2} & 0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} = J_2 = -i\pi(\sigma_2) \implies \pi(\sigma_2) = \begin{pmatrix}0 & \frac{1}{2}\\ -\frac{1}{2} & 0\end{pmatrix} \implies \pi(\sigma_2) = -\sigma_3.$$

$$-i\begin{pmatrix}\frac{i}{2} & 0\\ 0 & -\frac{i}{2}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = J_3 = -i\pi(\sigma_3) \implies \pi(\sigma_3) = \begin{pmatrix}\frac{i}{2} & 0\\ 0 & -\frac{i}{2}\end{pmatrix} \implies \pi(\sigma_3) = \sigma_1.$$

Thus, we have found how this representation acts on the basis elements of $Lie(SU(2))$, and so we have constructed a representation of $Lie(SU(2))$ using a highest weight of $\frac{1}{2}$. But what does this mean physically?

A particle in physics is defined as an irreducible representation, which our π is. So we have found a particle with a highest weight of $\frac{1}{2}$. We say that such a particle has a **spin** of $\frac{1}{2}$. Note that when the highest weight is $\frac{1}{2}$, the only possible weights are $\frac{1}{2}$ and $-\frac{1}{2}$. Physically, this means that our particle can either have spin up or spin down. We call these the **states** of the particle. So essentially, the spin of a particle corresponds to the highest weight, while the state of a particle can correspond to any of its weights. Next, we will discuss what it means to combine particles.

5.4 Combining Particles

Let π_1 and π_2 be two particles (irreducible representations) with spin (highest weight) $\frac{1}{2}$. If both π_1 and π_2 were in a state (weight) of $\frac{1}{2}$, then adding them together would give us 1. If one of them had a state of $-\frac{1}{2}$ and the other had a state of $\frac{1}{2}$, we would get 0. If they both had a state of $-\frac{1}{2}$, we would get -1. So when we combine these particles, they can end up with a state of 1, 0, or -1. But what does it mean to mathematically combine particles?

Consider, for example, a particle with 3 states and another particle with 2 states. If you combined these particles, you would get a total 6 different states in this new, combined particle (this is because $3 * 2 = 6$). Since we want a way to mathematically express multiplying particle states, what do we use? Well, since particles are irreducible representations, the way that one "multiplies" irreducible representations is through tensor products. For example, if A and B were particles, then $dim(A \otimes B) = dim(A)dim(B)$, and so it makes intuitive sense to say that combining particles is just like computing tensor products of irreducible representations, as we want a multiplicative-like operation.

For our particular case, we are looking at particles for $Lie(SU(2))$. For both π_1 and π_2, their corresponding weights are $\frac{1}{2}$ and $-\frac{1}{2}$. Since each of these particles have two possible weights, there are four possible ways to combine them. Let α_1 be the weight vector corresponding to a weight of $\frac{1}{2}$ for π_1 and β_1 be the weight vector corresponding to a weight of $-\frac{1}{2}$ for π_1. Similarly, let α_2 and β_2 be these corresponding weight vectors for π_2. Just like with our work in the previous sections, each of these particles will have J's that have the same properties as we used above. All of the J's will be labeled the same as before, with $J^{(1)}$ corresponding to π_1 and $J^{(2)}$ corresponding to π_2. If there is no superscript, then it is the J corresponding to the tensor product $\pi_1 \otimes \pi_2$.

Then we can say that $J_3(\alpha_1 \otimes \alpha_2) = J_3^{(1)}\alpha_1 \otimes I^{(2)}\alpha_2 + I^{(1)}\alpha_1 \otimes J_3^{(1)}\alpha_2$, since J_3 is a scalar multiple of a Lie algebra tensor product representation (where this follows by definition). This then equals $\frac{1}{2}\alpha_1 \otimes \alpha_2 + \alpha_1 \otimes \frac{1}{2}\alpha_2 = \frac{1}{2}(\alpha_1 \otimes \alpha_2) + \frac{1}{2}(\alpha_1 \otimes \alpha_2) = \alpha_1 \otimes \alpha_2$. Since this has a scalar multiple of 1, this weight vector $\alpha_1 \otimes \alpha_2$ has a corresponding weight of 1. Similarly, the weight vector $\beta_1 \otimes \beta_2$ has a corresponding weight of -1. When you combine an α with a β, in both cases you get a weight of 0. Thus, we can see how utilizing tensor products of these particles gives us states of 1, 0, or -1. Using mathematical language, this means that tensor products of these irreducible representations gives us weights of 1, 0, or -1. More could be done with tensor products of these representations, but we will stop our discussion here.

Chapter 6

Conclusion

We have reached the end of our journey through Lie theory. We started with some interesting knowledge from linear algebra and topology to get a common base for our audience. Then we looked at the interesting structures of Lie groups and Lie algebras, and their relationship. Next, we saw how representation theory could be utilized in this situation, in particular with our discussion of roots and weights. Finally, we used some of what we learned to take a brief trip through physics. Hopefully you learned some cool math (and physics) and enjoyed the ride!

Bibliography

[1] Ta Pei Cheng and Ling Fong Li. *Gauge theory of elementary particle physics*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1984.

[2] Bruce N. Cooperstein. *Advanced linear algebra*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, second edition, 2015.

[3] Jim Fowler. Tensor products, 2011. `https://www.youtube.com/watch?v=tpL95Sd7zT0`, visited 2021-03-22.

[4] Brian Hall. *Lie groups, Lie algebras, and representations*, volume 222 of *Graduate Texts in Mathematics*. Springer, Cham, second edition, 2015. An elementary introduction.

[5] R. Shankar. *Principles of quantum mechanics*. Springer, New York, second edition, 2008. Corrected reprint of the second (1994) edition.

[6] Benjamin Steinberg. *Representation theory of finite groups*. Universitext. Springer, New York, 2012. An introductory approach.

[7] Peter Woit. *Quantum theory, groups and representations*. Springer, Cham, 2017. An introduction.